



An Ontology to Support Behaviour Change Interventions: Barriers to Activity

Thesis submitted in accordance with the requirements of the University of Liverpool for
the degree of Doctor in Philosophy by

Yousef H. Alfaifi

October 2020

Dedication

To my family.

Acknowledgements

Many thanks should go to those who deserve it; first to my primary supervisor, Floriana, who provided me with indispensable advice and guidance in scientific research, for direction, for her expertise related to computer applications in health and health applications, and who facilitated all my administrative tasks within the computer science department. Thanks also go to my secondary supervisor, Valentina, who provided substantial and helpful advice during my studies, especially in the area of ontology. I would particularly like to thank her for her Semantic Web course that I attended, which reinforced many concepts, especially in web ontology language, and proved to be useful throughout my studies.

My profound gratitude is also expressed to my parents and siblings, for all their support in helping me to complete this study successfully. I also owe a great debt of gratitude to my wife and to my children, for standing beside me during my doctoral study. Without my family's support, we would not have been able to reach our goal.

Great thanks go to all my colleagues and friends, not only in Argumentation and Representation of Knowledge (ARK) group, but also across the entire computer science department, who were always helpful and assisted me whenever necessary. In addition, thanks also to all members staff in the computer science department, both academic staff and support staff, for providing a supportive research environment during my time there.

Finally, I express my huge appreciation for the generous financial support from the Saudi Arabian Government, represented by the Saudi Arabian Cultural Bureau in London, which provided me with the financial means to undertake my PhD in Liverpool.

Abstract

A complete view of health not only considers the presence or absence of disease, illness or disability, but also considers behavioural health. There are pharmacological interventions prescribed by specialists (e.g. doctors), that help the body resist a certain disease, such as inhalers for asthma sufferers. In a similar vein, there are behavioural interventions (non-pharmacological) that are aimed at influencing behaviour, with the goal to achieve and maintain a healthy lifestyle.

However behavioural interventions are often *ad hoc* and not well described, in terms of goals, actions and appropriateness to specific situations, and not always adhering to the CONSORT (Consolidated Standards of Reporting Trials) statement about standard in reporting and publishing interventions of behaviour change¹. To address this issue, a collaborative effort among 400 researchers, from different disciplines, and 12 Countries has resulted in the publication of a compendium of Behaviour Change Techniques (BCTs) [168]. The compendium includes 93 interventions to change human behaviour, with clear definitions and examples, at times including mechanisms of action and modes of delivery [165].

This thesis builds on the same effort, and expands it in two ways, one conceptual and one computational. From the conceptual point of view, we concentrate on a specific issue around successful BCTs, that is giving explicit consideration to *barriers* that might prevent the success of the intervention. From the computational point of view, we create a formal ontology of the barrier model, which can be used to support health behaviour applications in general, and BCTs in particular.

In order to give focus to the effort, we concentrate on barriers to physical activity for individuals affected by Type 2 diabetes (T2D). The work includes (i) identification of assumptions within the barriers domain knowledge and modelling of such barriers explicitly within the ontology; and (ii) extraction and communication of constructive insights such as advice (as suggestions for appropriate physical activity) to overcome barrier challenges. By following an established process of ontology development, this work constitutes also a walk-through case study in building an ontology from scratch, which expands on each phase of the process, including the evaluation stage.

¹<http://www.consort-statement.org>

Contents

Dedication	i
Acknowledgements	iii
Abstract	v
Contents	ix
List of Figures	xi
List of Tables	xiii
List of Algorithms	xiv
List of Abbreviations	xv
Publications	xvii
1 Introduction	1
1.1 Overview	1
1.2 Problem Statement and Motivation	2
1.3 Objective and Research Question	3
1.4 Contribution	4
1.5 Thesis Outline	5
1.6 Summary	6
2 Background: Models and Theories of Behaviour	7
2.1 Introduction	7
2.2 Understanding Behaviour	8
2.2.1 Behavioural Models	9
2.2.2 Theories of Behavioural Change	10
2.3 Behaviour Change and Maintenance	13

2.3.1	Behaviour Change Intervention	14
2.4	The Behaviour Change Techniques Taxonomy	16
2.5	Barriers and Determinants to Behaviour Change Interventions	18
2.6	Summary	20
3	Ontologies	21
3.1	Introduction	21
3.2	Defining an Ontology	22
3.3	Components of Ontologies	24
3.4	Types of Ontologies	25
3.5	Ontology Representation	27
3.5.1	Web Ontology Language	27
3.5.2	Description Logic	28
3.5.3	Resource Description Framework	29
3.6	Ontology Development Process	30
3.7	Ontology Evaluation	35
3.8	Ontologies in Behavioural Health	37
3.8.1	Behaviour Change Techniques as Hierarchical Taxonomy	37
3.8.2	Human Behaviour Taxonomy	39
3.8.3	Semantic Mining of Activity, Social and Health Data	39
3.8.4	The Neurobehavior Ontology	39
3.8.5	Health Behaviour Change Ontology	41
3.8.6	User Modelling Ontologies	41
3.8.7	Ontologies of Physical Activity Behaviour and Exercises	42
3.9	Summary	43
4	The Barrier Ontology Method and Development Process	44
4.1	Introduction	44
4.2	Requirements	45
4.3	Conceptual Barrier Model	47
4.3.1	Domain and Scoping	48
4.3.2	Systematic Review	49
4.3.3	Concepts of the Barrier Model	54
4.3.4	Relations of the Barrier Model	65
4.4	The Ontology Development Process	68
4.5	Summary	73
5	Evaluation	74
5.1	Introduction	74
5.2	Data-driven Approach	75
5.2.1	Comparison of the Barrier Ontology with Documents from the Domain	76

5.3	Competency Questions Technique	82
5.3.1	Survey	84
5.3.2	Analysis of Survey Results for Template Competency Question . . .	87
5.3.2.1	Template 1	89
5.3.2.2	Template 2	91
5.3.2.3	Template 3	92
5.3.2.4	Template 4	95
5.3.3	Conclusion of Competency Question Evaluation	96
5.4	Summary	97
6	Conclusion	99
6.1	Introduction	99
6.2	Summary of Thesis	99
6.3	Main Findings and Contributions	101
6.4	Limitations of the Work	103
6.4.1	Domain and Scoping Review of The Barrier Ontology	104
6.4.2	Evaluation Phase	104
6.5	Future work	105
6.5.1	Improving Maintenance Techniques	105
6.5.2	Expanding the Knowledge Base	106
6.5.3	Refinement of the Barrier ontology	106
6.5.4	Sourcing collaboration and publishing the ontology	107
6.6	Summary	108
	References	109
A	Systematic Review	138
B	Additional Details of the Barrier Ontology Evaluation	153
B.1	Introduction	153
B.2	Data-driven Approach	154
B.3	Competency Questions Technique	158
B.3.1	Survey Design and Content	158
B.3.2	Templates of Competency Questions	163
B.3.2.1	Template 1	163
B.3.2.2	Template 2	166
B.3.2.3	Template 3	169
B.3.2.4	Template 4	173

List of Figures

2.1	A model for understanding behaviour (the COM-B model) [170].	9
2.2	A model for understanding behaviour-change processes: Transtheoretic Model ² [213].	11
2.3	Theory of planned behaviour [3].	13
2.4	Components of the interventions, and relations among them [270].	15
3.1	Caption for LOF	34
3.2	A partial view of BCT Taxonomies.	38
3.3	Partial view from Human Behaviour Taxonomy [140, 199].	40
4.1	A high-level conceptual model of barriers to physical activity behaviour for T2D patients.	48
4.2	A flow-chart illustration of the Systematic Review.	51
4.3	Main barriers' classifications.	56
4.4	Hierarchical Health barrier classification (a continuation from Figure 4.3). .	56
4.5	Hierarchical Environmental barrier classification (a continuation from Fig. 4.3). .	57
4.6	Hierarchical Personal barrier classification (a continuation from Figure 4.3). .	57
4.7	Hierarchical Physical barrier classification (a continuation from Figure 4.3). .	58
4.8	Hierarchical Psychological barrier classification (a continuation from Figure 4.3). .	58
4.9	Hierarchical Social barrier classification (a continuation from Figure 4.3). .	58
4.10	Hierarchical classification of physical activity ontology (partial view).	59
4.11	Partial view from the disease ontology (identifying T2D as a disease) ³	61
4.12	Partial view from the General User Model Ontology (GUMO) [105].	63
4.13	Physical activity ontology imported into the conceptual barrier model.	70
5.1	Process diagram detailing the application of the data-driven approach in evaluating the Barrier Ontology.	77
5.2	Participation per profession distribution for the "Ontology to Support Behaviour Change Intervention: Barriers to Activity".	87
5.3	Bar chart depicting the agreement rate between the retrievable BO results and the SR for the CQs classified under Template 1.	90

5.4	Bar chart depicting the percentage similarity between the retrievable BO results and the SR for the CQs under Template 3.	92
5.5	Bar chart depicting the agreement rate between the retrievable BO results and the survey results for all CQs under Template 3.	94
5.6	Bar chart depicting the agreement rate between the retrievable BO results and the survey results for all CQs under Template 4.	96
B.1	A tree map depicting how domain experts classified barriers for CQs under Template 1. Each barrier classification was confirmed by a minimum of 33% of the total votes.	165
B.2	Bar chart indicating the category of activities that domain experts nominated as suitable for T2D patients who experience four specific barriers: ‘bad weather condition’, ‘lack of time’, ‘financial problem’ and ‘cultural norms . . .	175

List of Tables

4.1	A sample of the mapping of competency questions to functional requirements.	46
5.1	Some similarity between terms of the Barrier Ontology and the corpus . . .	80
5.2	A subset of each template's CQs, with their corresponding survey questions, target relations, and the ontological functional requirements they satisfy. Additional comments narrating how survey questions are derived from CQs are included as Meta-data. The complete table covering all functional requirements and survey questions is in Appendix (Table B.3).	85
5.3	Classifications assigned to barriers according to the Barrier Ontology (BO) results and Survey Result (SR) for CQs 1.1 and 1.2 under Template 1. ✓implies a barrier is identified as belonging to the corresponding Barrier Type, whereas X is the opposite. A match (Percentage Similarity (PS)) between the BO and SR column indicates that the ontology retrieved answers identical to those provided by experts in the survey.	89
5.4	Potential barriers according to the Barrier Ontology (BO) and Survey Results (SR) for scenarios 1 and 2 under Template 3. ✓implies a barrier is a hindrance to patient in response to the CQs, whereas X is the opposite.	91
5.5	Potential barriers to 'Football' activity according to Barrier Ontology (BO) and Survey results (SR). ✓implies a barrier is a hindrance to patient in response to the CQ 4.1, whereas X is the opposite.	93
5.6	Potential barriers to 'Swimming' activity according to Barrier Ontology (BO) and Survey results (SR). ✓implies a barrier is a hindrance to patient in response to the CQ 4.2, whereas X is the opposite.	93
5.7	Suggested activities according to the Barrier Ontology (BO) and Survey Results (SR) to limit bad weather condition barrier (CQ 4.1) and lack of time barrier (CQ 4.2) under Template 4. ✓implies an activity is a suggestion to patient in response to the CQs, whereas X is the opposite.	95
A.1	The main classification of identified barriers, with supporting source quotations	138
A.2	The main classification (Clas.) of identified barriers to physical activities (PAs), with supporting source quotations	141

A.3	Suggested physical activities (PAs) to limit identified barriers, with relevant source quotation	150
B.1	The complete similarity between terms of the Barrier Ontology and the corpus	154
B.2	Information about the websites crawled for information in generating the corpus.	156
B.3	A table of various CQs of each template, with their corresponding survey questions, target relations, and the ontological functional requirements they satisfy. Additional comments narrating how survey questions are derived from CQs are included as Meta-data.	159
B.4	Classifications assigned to barriers according to the Barrier Ontology (BO) results and Survey Result (SR) for CQs 1.3 and 1.4 under Template 1. ✓implies a barrier is identified as belonging to the corresponding Barrier Type, whereas X is the opposite. A match (Percentage Similarity (PS)) between the BO and SR column indicates that the ontology retrieved answers identical to those provided by experts in the survey.	166
B.5	Percentage volume of votes by domain experts, for activity preventing barriers for patients described in scenarios derived from Template 2 CQs.	169
B.6	Potential barriers according to the Barrier Ontology (BO) and Survey Results (SR) for scenarios 3 and 4 under Template 2. ✓implies a barrier is a hindrance to patient in response to the CQs, whereas X is the opposite.	169
B.7	Survey results for CQs under Template 3, ‘-’ means the barrier received less than 33% of votes from participants.	172
B.8	Potential barriers to <i>Yard work</i> activity according to Barrier Ontology (BO) and Survey results (SR).	172
B.9	Potential barriers to <i>Running</i> activity according to Barrier Ontology (BO) and Survey results (SR)	173
B.10	Suggested activities according to the Barrier Ontology (BO) and Survey Results (SR) to limit financial problem barrier (CQ 4.3) and cultural norms barrier (CQ 4.4) under Template 4. ✓implies an activity is a suggestion to patient in response to the CQs, whereas X is the opposite.	176

List of Algorithms

1	Search Wordnet to determine relatedness between ontology concepts and corpus terms.	80
---	---	----

List of Abbreviations

AI	Artificial Intelligence
BCT	behaviour Change Technique
BJGP	British Journal of General Practice
BNF	Backus-Naur Form
BO	Barrier Ontology
COM-B	Capability, Opportunity and Motivation-Behaviour
CONSORT	Consolidated Standards of Reporting Trials
CQ	Competency Questions
DL	Description Logic
DTP	Data Type Properties
FN	False Negatives
FP	False Positives
GO	Gene Ontology
GUMO	General User Model Ontology
HBCO	Health Behaviour Change Ontology
HBM	Health Belief Model
HTML	Hypertext Markup Language
ICD	International Classification of Diseases
ICF	International Classification of Functioning
IR	Information Retrieval
KB	World Health Organisation
KBS	Knowledge-Based Systems
LD	Levenshtein Distance
LDA	Latent Dirichlet Allocation
MeSH	Medical Subject Headings
NBO	Neurobehavior Ontology
NCBO	National Center for Biomedical Ontology
NCI	National Cancer Institute
NCIt	National Cancer Institute thesaurus
OBO	Open Biomedical Ontologies

ORBM	Ontology Restricted Boltzmann Machine
OWL	Web Ontology Language
RBM	Restricted Boltzmann Machine
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
SMASH	Semantic Mining of Activity, Social and Health Data
SR	Survey Results
T2D	Type 2 Diabetes
TIDieR	Intervention Description and Replication
TP	True Positives
TTM	Transtheoretical Model
UMs	User Models
UNO	User Navigation Ontology
W3C	World Wide Web Consortium
WHO	World Health Organisation
XML	Extensible Markup Language

Publications

Some contributions presented in this thesis have been published in peer-reviewed conferences, workshops and journals (abstract only). This thesis contains further developments, updates and details of these published works. All published articles which are directly related to this thesis are summarised below.

Conference papers

1. *Yousef Alfaifi, Floriana Grasso, Valentina Tamma: Towards an Ontology to Identify Barriers to Physical Activity for Type 2 Diabetes [10]. Digital Health Conference, London, 2-5 July 2017.*

This paper presents a proposed principle model towards an ontology to identify barriers to physical activity for T2D. This paper introduces a novel method to identify barriers based on their signs or characteristics (Section 4.3.4). Two steps of the ontology development process (determine the domain and scope of the barriers, enumerate the terms of the barriers) are included in this paper. The barriers' terms and barriers' classifications are acquired and categorized, respectively, based on relevant studies. The Unified Medical Language System⁴ and Electronic Health Records⁵ are the existing ontologies that are suggested to link with the concept of the barrier.

2. *Yousef, Alfaifi, Floriana Grasso and Valentina Tamma, An Ontology for Psychological Barriers to Support Behaviour Change [11], Digital Health Conference, Lyon, 23-26 April 2018.*

This paper is an extension of our previous work to find a way not only to identify and classify the barriers based on their signs (previous paper), but also how to limit these barriers by suggesting a suitable type of physical activity. An innovative conceptual model of barriers to physical activity is introduced, which supports health behaviour applications, such as Behaviour Change Technique (BCT). Related existing ontologies including the human disease ontology and physical activity ontology are re-used in this research to identify more health barriers, and to suggest a suitable activity,

⁴<https://www.nlm.nih.gov/research/umls/>

⁵<https://www.healthit.gov/faq/what-electronic-health-record-ehr>

respectively. Other associated concepts in this model are the stage of change, patient and belief concepts. Finally, the model links to BCT to support the BCT's goals. This conceptual model is created by using the development ontology process ('Ontology Development 101') (see Chapter 3 for more details).

Workshop papers

1. *Yousef, Alfaifi, Floriana Grasso and Valentina Tamma, Developing a Motivational System to Manage Physical Activity for Type 2 Diabetes [9], Workshop on Artificial Intelligence for Diabetes in ECAI, p.22.*

This paper presents a preliminary framework towards motivation of diabetic patients to perform regular physical activity. The rule-based expert system (which uses *if-then rules*) is used in this framework to identify accurate barriers to physical activity. Based on the psychological researcher's belief that the motivation of patients (e.g. diabetic patients) to change behaviour (e.g. physical activity) is more efficient than just treatment alone, persuasive technology is included in this framework. This persuasive technology is defined as "learning to automate behaviour change". The Fogg behaviour model, which includes motivation, ability, and prompt elements, is suggested to present the concept of the phase of the persuasive strategy. Even though this framework manages specific behaviour within a particular demographic group, it can be applied to control other behaviours, other chronic diseases, and other demographic groups.

Journal Abstract

1. *Alfaifi, Y., Grasso, F. and Tamma, V., Abstract Publication, Developing a Motivational System to Manage Physical Activity for T2D. Journal of Diabetes Science and Technology (JDST).*

The abstract of this paper is published in The Journal of Diabetes Science and Technology (JDST). This paper is summarized in the above paragraph.

Chapter 1

Introduction

1.1 Overview

Human behaviour is defined as “anything an individual does in response to internal or external events” [139]. The many different internal and external factors influencing behaviour make achieving and maintaining a change in behaviour a complex psychological problem, and this becomes particularly crucial for implementing public health interventions to encourage healthy behaviour in the population.

Various studies have attempted to understand the complexity of human behaviour, and how to successfully bring about change and subsequent maintenance of such change. These find important applications in the field of public health, which is increasingly at the centre of governments agendas. With the now ubiquitous digitalisation of society, and most importantly health, it is possible to implement interventions in a more systematic way, by reaching out to a greater number of individuals than it had been possible with traditional methods.

The work presented in this thesis contributes to the effort of digitalising public health interventions, by concentrating on the problem of representing the knowledge necessary for such interventions to be informed and principled.

This chapter serves as an introduction to the work presented here, and will start by outlining the problem and motivation, which will later be expanded in Chapter 2. It will state the main objective and research questions of this thesis, and will outline a summary of the main contributions of the work. This will be followed by a list of the published, peer reviewed output of the work presented here. An outline of the document will then be

presented, followed by a summary.

1.2 Problem Statement and Motivation

According to the World Health Organisation (WHO) [180], a person's general health is not just the absence of disease, disability, illness, etc., but also incorporates the individual's mental and behavioural health. Therefore, behavioural or mental health is no less important than physical health, and should not be ignored [127]. Many studies have attempted to model and understand behaviour, specifically health behaviour. Based on these studies, researchers working in social and psychological sciences, as well as in the health psychology, have made efforts to develop interventions to legally and ethically influence human behaviour, either to change or to maintain a previous behaviour change, where by intervention it is meant a "coordinated set of activities designed to change specified behaviour patterns" [170].

Change and maintenance of behaviour through intervention is unsurprisingly a complex psychological problem: is human behaviour is highly dynamic and is affected not only by volatile and often changing factors - which can be both *internal factors* (e.g. emotion and mood) and *external factors* (e.g. environment and culture) [21, 126], but also by the complexity of the same intervention chosen [140, 168, 169], as each intervention comprises different components or characteristics that interact with each other in a complex way [168, 181].

Such complexity has resulted in a very diverse style of reporting on various interventions [1, 8], with often unclear or inadequate descriptions of studies, and differing or incomplete definitions, therefore making it hard to come to a consensus on standard nomenclature or terminology for reporting efficient interventions, or evaluating the impact of interventions on behaviour change and maintenance [65, 139, 168].

This uncertainty of terms and lack of consensus on the reporting of terms has not been confined solely to theoretical behaviour studies, but it also negatively impacts on the task of evaluating the accuracy of practical efforts to change behaviour, in sharp contrast to other types of clinical studies, [90, 275], with a recent review reporting of about 71% of behavioural (non-pharmacological) interventions inaccurately described compared only with 33% of pharmacological interventions [165].

We will report in Chapter 2 of some efforts to propose a standard for generating and presenting interventions, such as CONSORT (Consolidated Standards of Reporting

Trials)¹ [179] or the TIDieR checklist [275], but the work presented in this thesis builds on the Behaviour Change Technique (BCT) taxonomy [140, 168], an international cooperative effort among 400 researchers from different disciplines and over 12 countries, which has resulted in the publication of a compendium of 93 behaviour change interventions with examples and definitions, organised into 16 groupings.

The work presented in this thesis builds on the BCT taxonomy, and expands it in two important directions: one conceptual and one computational. From a conceptual viewpoint, the BCT taxonomy, and in general other attempts to classify behavioural change techniques, do not account satisfactorily and explicitly for the *barriers* that prevent an intervention from partially or wholly achieving its goals, despite these form a main component of some behavioural models, as we will describe in Section 2.2.1.

From a computational viewpoint, while the BCT taxonomy constitutes a useful framework for the health practitioners, we were interested in a more formal effort, which could lay the foundations for building computer systems able to detect issues and plan digital interventions accordingly. Finding out which interventions are effective requires robust strategies to organise the large knowledge base, in order to allow aggregation, integration and comparison of results from different studies, thus aiding detection of the effective interventions of behaviour change and maintenance. We take therefore the existing hierarchical taxonomy of BCT as the essential first step in developing an “ontology”, as an explicit, machine-readable ‘specification of conceptualisation’ [98], for behavioural change systems, which is focused on the notion of barrier. To the best of our knowledge, there is currently no existing conceptual model or ontology of barriers to behaviour, therefore this study provides an original and important contribution to building intelligent systems supporting BCT and other health applications.

1.3 Objective and Research Question

As mentioned in the previous section, while the BCT taxonomy is an important step towards a conceptualisation of concepts around behaviour change, the focus is mainly on mechanisms of action and modes of delivery of interventions [165], and there is currently no satisfactory conceptualisation, nor taxonomy, that identifies, classifies, and categorises the barriers to behaviour, and the strategies to overcome them. Therefore, the primary

¹<http://www.consort-statement.org>

goal of this study is create a forma ontology of barriers, to support not only an expansion of the BCT, but also other health behaviour digital applications. We can summarise this study with the following central question:

“How can we use ontologies to formalise the notion of barriers to behaviour change and their underlying assumptions in a machine-readable format to support health informatics applications?”

In order to give focus and direction to our effort, we chose to concentrate on a specific context, by identifying a particular scenario, and a particular behaviour within that scenario which was likely to exhibit a barrier mechanism. In terms of scenario, and in line with our main purpose to consider digital health interventions, we identified Type 2 Diabetes (T2D) as a suitable one. This is a chronic disease, and is one of the most prevalent diseases around the world, accounting for around 95% of the cases of diabetes. More than 422 million people were living with diabetes in 2014, and this number is expected to reach 552 million by 2030 [235] and 592 million by 2035 [109].

Whilst diabetes is normally mentally associated with medication (tablets or insulin injections), a lesser known fact is that T2D can be effectively managed by adopting healthy lifestyles, for instance engaging into regular physical activity, eating healthy, and stop smoking. In particular, physical activity is probably less immediately associated with T2D management, and therefore likely to generate barriers: we therefore identified physical activity as the behaviour to consider within the scenario.

Therefore, this thesis will answer the above main question through the investigation of the following subsidiary questions:

1. How can we build an ontology of “barriers to physical activity for T2D patients”?
2. How can we demonstrate the use of a formal methodology, including the notion of ontology reuse, to objectively support the Barrier Ontology?
3. How can we use the Barrier Ontology to produce suitable recommendations of physical activities which take into account barriers to such physical activity from T2D patients?
4. How can we evaluate our methodology, and offer general suggestions for ontology developers?

1.4 Contribution

The main contribution of this thesis is a machine-processable ontology of barriers to physical activity, especially tailored to patients with T2D, and obtained through a refinement of

a taxonomy of terms acquired via a systematic review. This ontology was validated with a two-fold approach, consisting of a data-driven and a competency question (CQ) based evaluation.

This main contribution leads to several sub-contributions:

1. A taxonomy of generic barriers to behavioural change, which provides the foundation for the ontology.
2. The identification of assumptions within the barrier's domain knowledge, and their explicit modelling within the ontology.
3. A comprehensive analysis that annotates relations between barriers and physical activities for T2D patients, enabling researchers and software developers to gain a greater understanding of the underlying information structure.
4. The introduction of a hybrid re-usable framework for building and evaluating an ontology driven from data.

1.5 Thesis Outline

The rest of the thesis is organised as follows:

Chapter 2: Background: Models and Theories of Behaviour. This chapter presents the psychological background to the study in the thesis, by reviewing studies on human behaviour explaining how it can be changed and maintained, and how interventions can be build to influence behavioural change.

Chapter 3: Ontology. This chapter unpacks intrinsic details pertaining to the subject of ontologies. The chapter commences with a discussion of different ontology definitions, covering both the computing and philosophical perspectives. This is followed by a review of ontology representation techniques. Additionally, the chapter includes a narration of the ontology development process and ontology evaluation methods (data-driven and competency question based) implemented in the research study, which then Chapters 4 and 5 discuss when evaluating the Barrier Ontology. Finally, this chapter complete the previous one by surveying efforts to create ontologies or taxonomies for behaviour and behaviour change.

Chapter 4: The Barrier Ontology Method and Development Process. This chapter breaks down the sequential steps followed in the development of the thesis' main contribution: The Barrier Ontology. It begins with a clear articulation of the ontology's requirements, then it presents the systematic review of literature used to extract terms and

relations. A high-level conceptual barrier model is built by mapping the various concepts or classes entailed in the taxonomy to each other depicting the identified relations.

Chapter 5: Evaluation. This chapter discusses how we apply two different approaches in evaluating the Barrier Ontology: Firstly, a data-driven approach is used to compare the ontology's vocabulary of terms with a corpus of topics, summarising content in randomly extracted web articles on the domain. The comparison computes measures including precision, recall, and harmonic mean, which reveal the extent of semantic and syntactical similarity between the ontology and the information of the domain it represents. The second approach involves formulating a set of "Competency Questions" and extracting answers to the questions from the Ontology. A validation of these questions is presented which consists of a case study with human domain expert participants through a survey. Percentage Similarity is used to determine the similarity between two sets of answers, hence estimating the likelihood that the ontology would achieve its purpose successfully.

Chapter 6: Conclusion. This chapter summarises the answers to the research questions presented in this Chapter 1, with reference to the specific chapters where these were discussed. This is followed by summaries of the main findings and contributions, and of the limitations of the approach. Finally, proposed future work to scale and enhance the contributions made is discussed.

1.6 Summary

This Chapter frames the problem investigated in this thesis, by presenting the main research questions, providing the motivations which drove the work, grounded in efforts in behavioural psychology to design effective interventions for behaviour change, and identifies the main contributions of this thesis. A map of the rest of the document is also presented. Chapter 2 will now move to a literature review of behaviour and behaviour change, providing more detail on what has been introduced in this chapter.

Chapter 2

Background: Models and Theories of Behaviour

2.1 Introduction

In Chapter 1, we presented a general overview, problem statement, motivation of the thesis, and summary of the contributions and the research questions in this thesis. Before moving to the main contribution of this thesis, that is the process of creating an ontology of barriers to physical activity for Type 2 Diabetes (T2D), we will discuss in this chapter relevant background to the area of behaviour change and maintenance, in order to better motivate our choice of focusing on barriers to behaviour. A more specific review on efforts towards the design of digital systems promoting behaviour change using an ontology is deferred to Section 3.8.

This chapter will discuss, in Section 2.2, human behaviour and its key aspects, particularly exploring behaviour models and behavioural change theories, where the former uniquely dissect factors underlying and influencing behaviour, and the latter facilitate developing interventions aimed at changing or influencing human behaviour. Interventions that can be used to change and maintain behaviour are discussed in detail in Section 2.3. The Behaviour Change Technique (BCT) standard is presented in Section 2.4, with a discussion on barriers and determinants in Section 2.5. Finally, the chapter is summarised in Section 2.6.

2.2 Understanding Behaviour

Human behaviour encompasses mental aspects, physical aspects, and all the ways humans act and interact with each other. Based on behavioural studies, human behaviour is defined as “anything an individual does in response to internal or external events” [139]. This means human behaviour is highly complex and dynamic; it is influenced by many ever-changing internal factors (e.g. emotion and genetic make-up) and external factors (e.g. culture and the environment) [87, 126, 140]. It is often useful to characterise, rather than behaviour in general, behaviours within a specific context or scenario, so we can discuss for instance, among others, dietary behaviour, physical activity behaviour, smoking behaviour, shopping behaviour, or adherence behaviour [161]. Most of the studies presented in this chapter are related to the area of public health and medicine, not only because we also discuss health-related behaviour, but also because most behaviour studies are associated with this area [55], however we will also discuss more general models (e.g. Fogg behaviour model [82]), as they can be applied to any particular behaviour, including health behaviour.

There are many different theories and models related to behaviour and behaviour change interventions. *A. Darnton* [61] acknowledges that models of behaviour and theories of behaviour change have significant overlaps, however also highlights and extensively discusses their subtle differences. *Darnton* goes on to provide an intuitive distinction between the two, indicating that models of behaviour (behavioural models) are helpful in understanding specific behaviours by identifying the underlying factors which influence them, whereas theories of behavioural change show how behaviours change over time and how they can be purposefully changed.

Behavioural models are largely informed by theories which diagnostically explain factors underlying or influencing behaviour. Examples of behaviour models include the Capability, Opportunity and Motivation-Behaviour (COM-B) Model [170] and the Health Belief Model (HBM) [119]. On the other hand, theories of behavioural change are focussed on supporting interventions to change current behaviour or to facilitate adoption of new behaviour. Examples of theories include the theory of behavioural change [19], the Theory of Planned Behaviour (TPB) [3] and the Transtheoretical Model (TTM) [213] or theory [61, 239].

Some behavioural models and theories of behavioural change, including those aforementioned above, are discussed below.

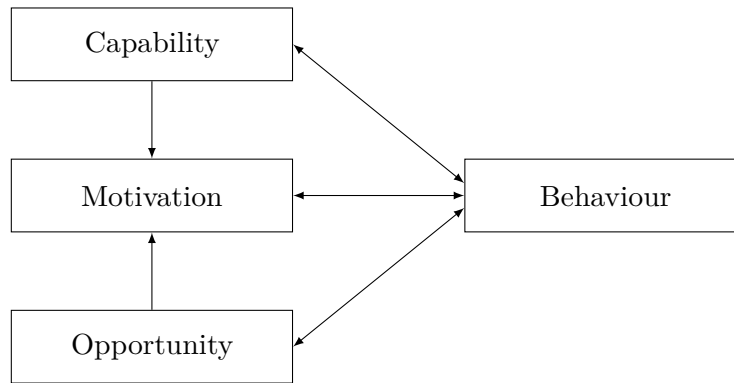


Figure 2.1: A model for understanding behaviour (the COM-B model) [170].

2.2.1 Behavioural Models

We now present two prominent examples of behaviour models: COM-B and HBM.

Capability, Opportunity and Motivation-Behaviour Model

The Capability, Opportunity and Motivation-Behaviour (COM-B) Model or system [170] is a recent widely cited model published in 2011 by Susan Michie *et al.* [164, 170], and subsequently expanded to the compendium of Behaviour Change Techniques (BCTs) [168] mentioned earlier, and discussed in detail in Sections 2.4 and then 3.8.1.

The components of COM-B give the essential conditions for a particular behaviour to occur, interacting in a complex way to generate behaviour. Such components are *Capability*, *opportunity*, and *motivation*. *Capability* is defined as the physical ability (skills) and mental ability (knowledge) required to perform a particular behaviour. *Motivation* involves all psychological brain processes, such as decision-making and emotion, to guide and prompt behaviour. *Opportunity* concerns external factors that support the behaviours occurrence, such as environmental factors. Figure 2.1 shows the components of the behaviour system; the single and double arrows represent the interactions between components, and interactions with behaviour, respectively. Opportunity and capability can influence motivation, while the behaviour can be altered or affected by all three components (as shown in Figure 2.1) [170].

These components are common to other studies, even if they might be named differently, for example, Fogg’s behaviour model uses ability, trigger and motivation [82], as opposed to capability, opportunity, and motivation, respectively [170]: a common insight is that exerting influence on the behaviour’s components (e.g. motivation) supports the task of

changing behaviour and maintaining the change (Section 2.3) [82, 170].

Health Belief Model

The Health Belief Model HBM [119] is one of the oldest behaviour models and the most frequently used model in psychological health. The HBM was developed in 1950 by social psychologists at the U.S. Public Health Service [234]. The primary goal of this model is to understand behaviour surrounding individuals who ignore medical advice to prevent disease, and who reject regular check-ups to detect diseases early [119]. The HBM is based on psychological and behavioural models, such as the belief that the real wish to avoid disease, and the belief that a specific action, either physical or mental, prevents or cures the disease, are related to health behaviour. The HBM includes six constructs: perceived susceptibility, perceived severity, perceived benefits, perceived barriers, cue to action and self-efficacy. The last two constructs (cue to action and self-efficacy) were not present in the original model, but were added later by psychologists [119].

2.2.2 Theories of Behavioural Change

Our underlying aim is to contribute to the efforts implementing successful behaviour change interventions. In order to be effective, implementations of behaviour change interventions need to be based on knowledge on how behaviour can be affected [168, 170], and also on understanding how these studies are to be interpreted [8, 43, 188]. Despite this, a survey looking at 235 empirical applications of interventions has shown that as many as 77.5% of these ignore the theoretical foundation [64, 65]. The natural result of this is an ineffective implementation and flawed study of behaviour change interventions [167]. Therefore, to produce reliable and dependable interventions, the existing theoretical studies - consisting of theories, models, systematic literature reviews, analytical and experimental studies must be taken into account [65, 139, 168]. This helps to explain the importance of the theoretical knowledge when looking to build a contribution that is dependable and practical, and this approach is what has been adopted in this research.

Two prominent theories of behavioural change, TTM and TPB are presented below.

Transtheoretical Model

The Transtheoretical Model (TTM) [213] or theory [61, 239] was developed in the late 1970s by Prochaska and DiClemente [213]. Based on this model, the change of behaviour is the

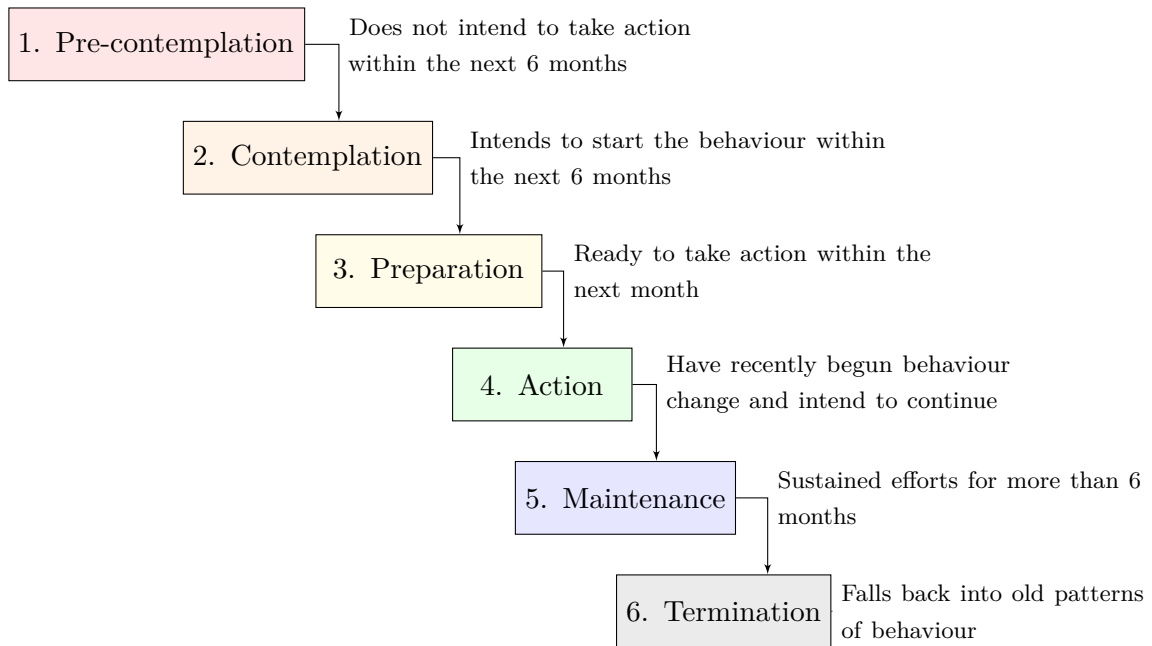


Figure 2.2: A model for understanding behaviour-change processes: Transtheoretic Model¹ [213].

process of progressing through stages of the change cycle. Because of this change cycle, the TTM is often referred to as the stage of change model [212]. This is despite these stages being only one of several components of the TTM, such as self-efficacy, levels of change, and processes of change. A particular behaviour (smoking) is used in the earlier publications to exemplify the behaviour change problem, but the model was applied to other behaviours.

The TTM change cycle is described in six stages (Figure 2.2) and estimates the period of time needed for a transition from one stage to the next, or the duration of each stage, in six-month increments. The six stages are *pre-contemplation*, *contemplation*, *preparation*, *action*, *maintenance*, and *termination* [212, 213]:

1. *Pre-contemplation* (Not Ready): An individual in this stage either does not intend to take action to change the behaviour, or is not thinking about behaviour change. This stage is “usually measured as the next six months”.
2. *Contemplation* (Getting Ready): An individual in this stage is starting to think about changing the behaviour or is starting to think about taking action. This stage is

¹This figure is not provided in [213].

measured “in the next six months” in order to observe the individual’s intention to change.

3. *Preparation*: An individual in this stage is preparing to take action (e.g. making preparations to start regular physical activity) in the near future, which “usually measured as the next month”.
4. *Action*: An individual in this stage is taking action and working towards maintaining the desired behaviour. To evaluate the individual’s performance in this stage, the specialists need to review individual’s action “within the past six months”.
5. *Maintenance*: An individual in this stage is maintaining and keeping the desired behaviour (e.g. regular physical activity maintained for more than six months).
6. *Termination*: This is the last stage of the change process, where an individual has completed the change cycle, and has zero temptation to relapse (100% self-efficacy).

The TTM is important also as it has been often adopted to produce digital interventions, and in particular it was used to model behaviour and behaviour change within existing behaviour change ontologies, such as the Health Behaviour Change Ontology (HBCO) (Section 3.8.5), which we also utilise, as detailed in Section 4.3.3.

Theory of Planned Behaviour

The Theory of Planned Behaviour (TPB), published in 1991 by *I. Ajzen* [3], as an extension of the theory of reasoned action by the same author [4]. The TPB aims to predict an individual’s intention to engage in a behaviour at a specific time and place. The components within the TPB are *behavioural intention*, *subjective norms*, *attitude toward the behaviour*, *perceived behavioural control*, and *target behaviour*. The *behaviour intention* captures motivational factors that influence a specific behaviour, such as how much effort someone is willing to put into achieving the target behaviour. When the intention is stronger, it stands to reason that the possibility of the behaviour occurring is increased. The *subjective norm* component refers to socially accepted conventions and socially imposed pressures upon an individual, which restrict or encourage the performing of particular behaviours. The *perceived behavioural control* component refers to an individuals perception of ease or difficulty in performing a given behaviour. The *attitude toward the behaviour* component captures an individuals personal evaluation of a particular behaviour, being either favourable or unfavourable. Finally, the *target behaviour* component captures the behaviour that is intended to be changed or reinforced. Figure 2.3 presents these components and the

relations among them.

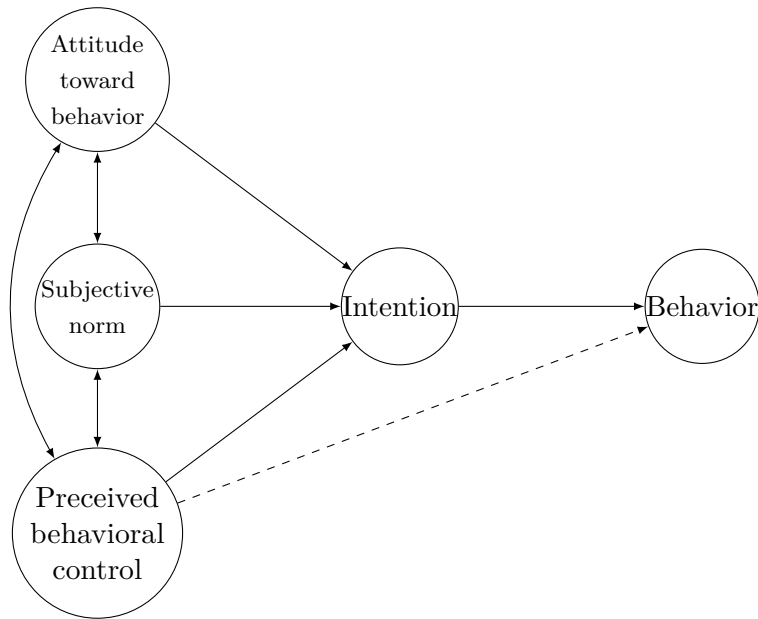


Figure 2.3: Theory of planned behaviour [3].

The formulation of TPB makes it particularly useful in predicting why people adopt some behaviours and what are the perceived barriers to such adoption, in many cases in a more efficient way than other models such as the HBM [36, 187, 216], which makes this an important piece of research to consider in our study.

2.3 Behaviour Change and Maintenance

Behavioural science is involved in studying and understanding the complexity of human behaviour, not only to change undesirable behaviour to desirable behaviour, but also to maintain the change on a long term basis, as there is no great benefit to a short term behaviour change [139].

Behaviour change is defined as “a set of interventions and strategies designed to change specified behaviour patterns” [79, 170] to transform or modify of human behaviour. Interventions are defined as “as coordinated sets of activities designed to change specified behaviour patterns” [170]. Observing and evaluating any behaviour change, however small, takes a long time [81, 121, 211], however, regardless of the amount of change in such

behaviour, it is still important to mark a positive impact on the individual, as well as a positive impact on the community and society as a whole (e.g. public health) [8, 112, 139]. For example, individuals reducing the number of cigarettes they smoke, and increasing their level of physical activity, are enacting a positive behaviour change which will have a positive effect on their life [139].

Behaviour change maintenance is defined as “The continuous performance of a behaviour following an initial intentional change” [139]. In other words, maintaining the behaviour reinforces the current behaviour. In models such as the Transtheoretical model, behaviour maintenance usually occurs at advanced stages (action and maintenance stage) (Section 2.2.2) [213]. This is one of the reasons that there is no separation or difference between behaviour change and behaviour maintenance in some behaviour studies, as maintenance of behaviour is included alongside the behaviour change. An individual continuing to not smoke or resisting the urge to smoke again [69], and performing regular physical activity for an hour a day [112] are both examples of behaviour maintenance.

Comparison of the effective interventions to maintain the behaviour and the effective interventions to change the behaviour shows that there are less evidence-based interventions surrounding behaviour maintenance than there are surrounding behaviour change [65, 275]. This is due to a lack of studies evaluating interventions on a long-term basis. Additionally, the impact of any intervention may weaken over time [57, 71, 139].

An understanding of behaviour change and maintenance alongside the complexity of behaviour is required to decide on an effective intervention that has a positive influence on behaviour change and maintenance [139, 161, 168, 275].

2.3.1 Behaviour Change Intervention

In order to change behaviour and maintain this change, efficient interventions are implemented by behavioural researchers. An intervention (sometimes referred to as technique) of behaviour change is defined as “as coordinated sets of activities designed to change specified behaviour patterns” [164]. ‘*Feedback*’, ‘*suggestion or advice*’, ‘*argument*’ and ‘*reinforcement*’ are examples of interventions [168]. To implement and understand these interventions, their characteristics or components have to be explained. The interventions include some characteristics which are defined by Davidson *et al.* [62] as “who delivers the intervention, to whom, how often, for how long, in what format, in what context, and with what content”. From this definition, there are six main components of interventions: *intervention* (content

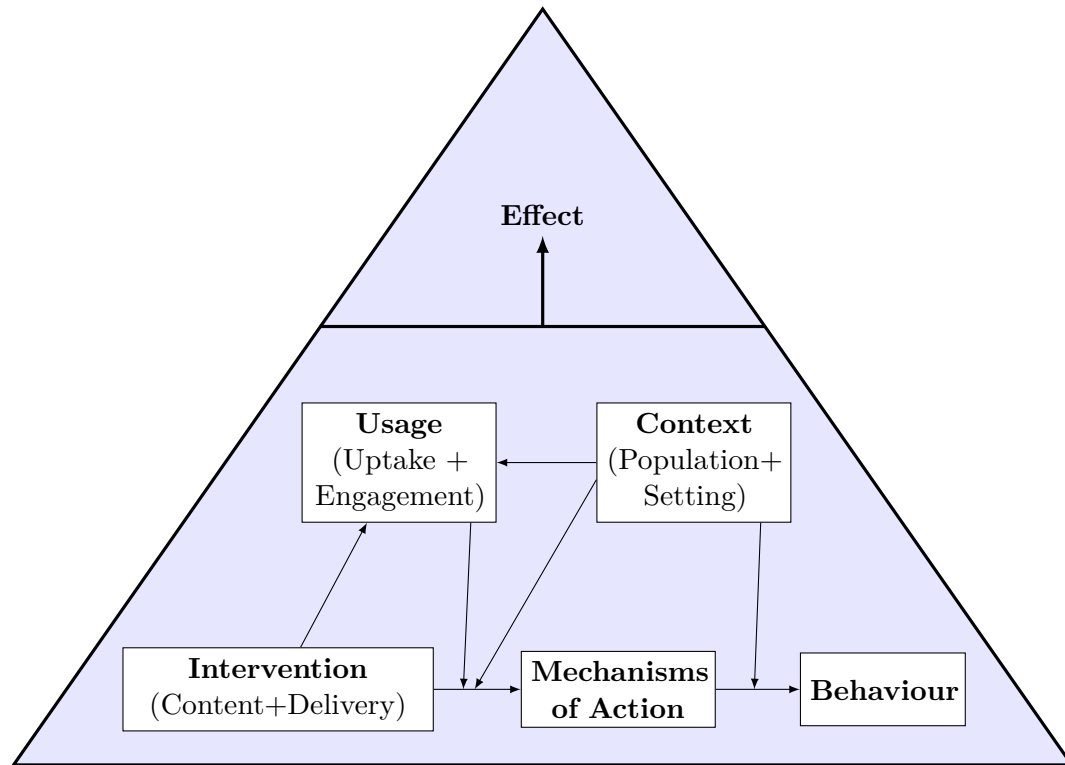


Figure 2.4: Components of the interventions, and relations among them [270].

and delivery), *mechanisms of action*, *behaviour (target behaviour)*, *context* (population and setting), *intervention usage (uptake and engagement)*, and *intervention effect* [270]. The components of interventions and the relations between them are shown in Figure 2.4.

The first and main component is *intervention* itself, which can be observed through the delivery content (e.g. GP advice or coach feedback) and through the delivery method (e.g. meeting, SMS, or mobile application). The second component is the *target behaviour*, which is the specification of the behavioural outcome; this comprises the target behaviour to be changed (e.g. smoking) or maintained (e.g. not smoking) [53, 275], and how the change and maintenance is to be assessed after the intervention [69, 147]. The third component is *mechanism of action*, which is defined as “the processes through which a behaviour change technique affects behaviour” [165]. It aims to represent the relationship between the targeted behaviour and the type of interventions. For example, emphasis is given towards health risks (increased diabetes complications) and towards psychological risks [68] (increased anxiety) to encourage the target person or group to perform or maintain the

physical activity [207]. The mechanism of action is surrounded by debate and controversy in the theoretical knowledge of behaviour and behaviour change community. This debate is centred on the need for a consensus to be established about how the mechanisms of action, that are specified in the theoretical work, are able to be mapped or linked to specific intervention components. That is to say that a deeper understanding is required of how the theoretical mechanisms of action interact with individual intervention components [165]. The fourth component is the *interventions usage*, which is known as the acceptance (uptake) and the involvement (engagement) of the intervention of the target person or group. The fifth component is *context* of the intervention, which contains the characteristics of the target person or group (target population) and other related external factors (setting). The characteristics refer to the health and psychological condition or status of the target person or group. The setting includes social factors (e.g. culture, education, social and ethnic customs) and environmental factors (economic status, time restriction and location). The last component is *effect*, which measures the result (outcome) of a particular component, such as interventions and usage, on the targeted behaviour. Figure 2.4, which is published by West and Michie [270] demonstrates the components of behaviour change intervention, the relationships between them, and how they interact.

The complexity of these components (e.g. mechanisms of action), as shown in Figure 2.4, must be taken into account during the development of behaviour change applications, using concepts such as an ontology (Chapter 3). When, studying and understanding behaviour (e.g. context), behaviour change (e.g. intervention), the structure of human behaviour and the cycle of behaviour change and maintenance, it is necessary to develop efficient interventions based on scientific evidence [169].

The goal of behavioural medicine is to understand the mechanisms of action, the processes of behaviour change, etc. in order to decide on suitable interventions to change the target behaviour and to then maintain the change. This understanding is acquired from the theoretical knowledge of behaviour and behaviour change [168, 169].

2.4 The Behaviour Change Techniques Taxonomy

The inherent complexity of interventions to influence behaviour, as mentioned in Section 2.3.1 and demonstrated in Figure 2.4, has at times resulted in the poor descriptions and definitions of interventions across behaviour studies. While there is an established practice in reporting and publishing clinical interventions in an evidence-based manner, this is not always true

for behavioural ones. For example, CONSORT (Consolidated Standards Reporting Trials)² is “intended to improve the reporting of a randomized controlled trials (RCTs), enabling readers to understand a trial’s conduct and to assess the validity of its results” [179], and evidence suggests that using it increases the quality of reports of RCTs. It provides guidance for researchers to publish and report eHealth and mHealth interventions [76]. CONSORT is published in Dutch, English, French, German, Japanese and Spanish, and it includes a precise description of the content, delivery method and mechanisms of action related to the interventions [275]. In fact, a systematic review [155] showed that 67% of pharmacological interventions from medical specialists were accurately defined and described. This is in sharp contrast with non-pharmacological interventions, of which only 29% were accurately defined and described. Even when using standard reporting, mechanisms like CONSORT do not provide details pertaining to “who delivers the intervention, to whom, how often, for how long, in what format, and in what context, and with what content” [168]. Other checklists exist, such as the Template for Intervention Description and Replication (TIDieR) which defines the primary elements (i.e. interventions’ characteristics and content, mode of delivery, target behaviour, context, uptake and engagement, and intervention effect) [160] (Figure 2.4), that must be reported for any intervention, but includes only the procedures of delivery (mode of delivery) of interventions (such as face-to-face) instead of the intervention’s content (e.g. goal setting) [275].

To bridge this gap, an effort by 400 researchers and specialists from various disciplines, including behavioural medicine, from 40 countries, collaborated to create the "Behavior Change Technique Taxonomy" [168, 276] in which a behaviour change technique (BCT) is defined as “an observable, replicable, and irreducible component of an intervention designed to alter or redirect causal processes that regulate behaviour; that is, a technique is proposed to be an *active ingredient*” [168]. Since then, this has become a unified method to publish and report evidence-based behaviour change interventions. The taxonomy includes 93 interventions with clear definitions and examples. It covers many different behaviour types, such as healthy eating, physical activity [162], smoking cessation [271], and changing professional behaviour [117, 168].

The organisation of BCTs in a taxonomy is particularly useful to our purposes as the first approximation to a more complex knowledge representation in form of an ontology, as we will discuss in Chapter 3.

²<http://www.consort-statement.org>

2.5 Barriers and Determinants to Behaviour Change Interventions

In models such as the COM-B model (Section 2.2.1) it is postulated that in order for a particular behaviour change to occur and be maintained, some behaviour's components should be influenced [82, 170]. This influence on these components have to come together at the same time to achieve the goal of modifying behaviour and maintaining the behaviour change, otherwise the influence on behaviour does not happen [82].

The BCT taxonomy also links directly to different behavioural determinants to achieve the target behaviour [154, 166]. These determinants are defined as "the factors that require change for the goal behaviour to occur" [154].

Theories and models of behaviour help identify these determinants, such as the already mentioned Health Belief Model [119], Bandura's theory of behavioural change [19], or the Theory of Planned Behaviour (TPB) [3], or more recently the Behavioural Reasoning Theory [272].

In [78] Fishbein *et al* note that for a person to engage into a behaviour, the following circumstance need, all or in part, to apply:

1. *The person has formed a strong positive intention (or made a commitment) to perform the behavior.*
2. *There are no environmental constraints that make it impossible for the behavior to occur.*
3. *The person has the skills necessary to perform the behavior.*
4. *The person believes that the advantages (benefits, anticipated positive outcomes) of performing the behavior outweigh the disadvantages (costs, anticipated negative outcomes); in other words, the person has a positive attitude toward performing the behavior.*
5. *The person perceives more social (normative) pressure to perform the behavior than to not perform the behavior.*
6. *The person perceives that performance of the behavior is more consistent than inconsistent with his or her self-image, or that its performance does not violate personal standards that activate negative self-sanctions.*
7. *The person's emotional reaction to performing the behavior is more positive than negative.*
8. *The person perceives that he or she has the capabilities to perform the behavior under a*

number of different circumstances; in other words, the person has perceived self-efficacy to execute the behavior in question.

Factors and determinants can be enablers (they facilitate a behaviour) or barriers (they prevent a behaviour), and both sets of factors need to be addressed in order for a change to occur. In an example such as physical activity behaviour, intention, self-efficacy and barriers (e.g. lack of time and tiredness), may be the factors that require intervention [80, 86].

In our study, we concentrate on the second set of factors, and we focus our attention to the barriers to a behaviour. Barriers are often not addressed to a sufficient extent, with behaviour change campaigns more likely to concentrate on enablers. However, McKenzie-Mohr [158] notes for instance that interventions aimed at "enhancing knowledge and creating supportive attitudes" often have very little impact, as the complexity of adopting a behaviour is not like changing consumer preference, while *"[i]n contrast, promoting engagement in a new activity, such as walking or biking to work, is much more complex. An array of barriers to these activities exist, such as concerns over time, safety, weather, and convenience. The diversity of barriers that exist for any sustainable activity means that information campaigns alone will rarely bring about behaviour change."* [158].

Different barriers prevent interventions from achieving their goal of changing and maintaining a specific behaviour [163]. The term 'barriers' first appeared in 1974 in Baker's health belief model [226] as "the obstacles or impediments to taking action to reduce the threat of illness" [157]. The barriers were also included in different behaviour models such as the HBM, defined as "perceptions concerning the unavailability, inconvenience, expense, difficulty, or time consuming nature of a particular action" [157, 204]. Barriers can be categorised into different types, depending on whether they refer to health, physical, personal, environmental, psychological, or social factors. The identification of barrier to behaviour in particular for patients with T2D was highlighted as crucial in behaviour change interventions [222], which confirms the appropriateness of the scenario we use in this work. Addressing barriers in a conceptual model for interventions will support not only the digitalisation of behavioural change techniques, but also other health behaviour applications, as appropriate patient advice to prevent or overcome these barriers will give more efficacy to the behaviour change [163].

2.6 Summary

Comprehensive health assessment includes both physical and behavioural health. Despite the complexity of human behaviour, which can be affected by many internal and external factors, there have been clear efforts to create interventions to improve behaviour of patients, to swap unhealthy behaviour for healthier behaviour. Theories and models from the behaviour research area support the understanding of human behaviour before, helping researchers deploy interventions to change and maintain behaviour. Having reviewed the most influential theories and models of behaviour, we have identified the factors that need to be taken into account to implement behaviour change interventions, and we have noted that barriers, as factors preventing change, are both not sufficiently addressed and extremely crucial for long term, complex behaviour change. We can now move to address the conceptualisation of these factors into an ontology which can support digital behavioural change interventions, and in Chapter 3 we begin by explaining and reviewing the field of ontology engineering, before reviewing applications within the behaviour and behaviour change domain that make use of ontologies.

Chapter 3

Ontologies

3.1 Introduction

In Chapter 2, we reviewed human behaviour, behavioural models, to understand human behaviour and behaviour change. This chapter provides a review of research on ontologies, before moving to report on the main contribution of this thesis, the Barrier Ontology.

Ontology is a term that originates in Philosophy, before being borrowed by Computer Science, and Knowledge Engineering. In Section 3.2, we define what is meant by ontology in Philosophy; one of the most accepted definitions of an ontology in Computer Science was published by Gruber in 1993: “*an explicit specification of conceptualisation*” [97]. This chapter discusses the refinements that have been made to Gruber’s definition by others, and how these relate to the structure, creation and evaluation of ontologies in computer science; from here we refer to ontology in computer science simply as ontology.

Sections 3.3 and 3.4 present components and types of an ontology, respectively, highlighting those used in this thesis. Formal languages for representing an ontology are discussed further together with ontology representation (Section 3.5). Section 3.6 discusses the ontology development process, focusing on the “Ontology Development 101” methodology, which is the one followed in this thesis to develop the Barrier Ontology.

Evaluation of an ontology is one of the most debated topics in the ontology field. There is no set approach to ontology evaluation. Several proposed approaches to evaluate an ontology are discussed in Section 3.7. Finally, Section 3.8 presents some taxonomies and ontology efforts that are related to behaviour and behaviour change. The chapter is summarised in Section 3.9.

3.2 Defining an Ontology

The term *ontology* appears in both Philosophy and in Computer Science. Philosophically, an ontology is defined as “the branch of metaphysics dealing with the nature of being” [13, 260] or the “theory of existence” [177]. This concept has been successfully imported in Computer Science, in the area of research dealing with knowledge representation, and several new definitions were produced, which we list here in order of their published date i.e. from *oldest to newest*.

Neches *et al.* “*An ontology is the basic terms and relations comprising the vocabulary of a topic area as well as the rules for combining terms and relations to define extensions to the vocabulary*” [186]. This definition, published in 1991 by Neches *et al.*, is to our knowledge the first definition of ontology in Computer Science [92, 93, 195]. In addition to the given definition, the authors detail what they consider an ontology’s main components, that is terms and relations, and how to combine these by using rules to build an ontology [92, 93]. Interestingly, this definition indicates that terms include both explicit terms and terms inferred by using rules [93].

Gruber. “*An ontology is an explicit specification of a conceptualization*” [97]. This is probably the best known definition, and is very flexible in that an ontology can take one of multiple different forms; an ontology can be seen as a control vocabulary (collection of terms), a glossary (a list of terms and their meanings), or a set of taxonomies (terms with their relations) [156, 189]. An ontology may also be a set of definitions of classes, properties, and constraints on the way those classes and properties can be employed. As a minimum, an ontology includes a taxonomy, that is “a hierarchical is-a relation between concepts” [37]. When we consider the ontology development process, we will see that defining *the classes and the class hierarchy (taxonomy)* is the main step (step 4) of the process (Section 3.6).

The above definition of an ontology, was published in 1993 by Gruber. Despite coming 2 years after the one by Neches, Gruber’s definition has become the most accepted and the most referenced definition of an ontology in the literature [92, 93], hence, most definitions of an ontology today, are refinements of Gruber’s definition [93, 243]. We include some examples of these below.

- **Borst.** “*Ontologies are defined as a formal specification of a shared conceptualization*” [35].

The definition by Borst is a formal specification, covering the nature and purpose of an ontology. The formal specification helps it to be machine readable, and the reference to a shared conceptualisation looks towards offering mediation between a person and a system. This definition is distinguished from Gruber's by the term "shared", emphasising that the ontology needs to be agreed upon by the stakeholders [210].

- **Studer *et al.*** *"An ontology is a conceptualization that refers to an abstract model of some phenomenon in the world by having identified the relevant concepts of that phenomenon"* [243].

This definition by Studer *et al.* combines the definitions by Gruber and Borst, giving a more specific explanation of the two.

- **Swartout *et al.*** *"An ontology is a hierarchically structured set of terms for describing a domain that can be used as a skeletal foundation for a knowledge base"* [247].

The definition by Swartout *et al.* uses the fact that a domain-specific ontology comprises a set of related terms arranged in a hierarchical structure for a specific domain. Heuristics are then used to prune the given ontology when needed [92].

- **Bernaras *et al.*** *"An ontology provides the means for describing explicitly the conceptualization behind the knowledge represented in a knowledge base"* [28].

From Bernaras *et al.*'s perspective, an ontology is an initial construction of the knowledge base. The ontology is then extended, refined and augmented to include more definitions when new applications for the ontology are created.

A few keywords resonate across most of the definitions narrated above:

1. *Explicit*; implying that the type of concepts used, and the constraints on their use are clearly expressed [243].
2. *Formal*; "referring to the fact that the ontology should be machine-readable" [243].
3. *Shared*; "reflects the notion that an ontology captures consensual knowledge, that is, it is not private to some individual, but accepted by a group" [243].
4. *Specification*; a formal description of how something is to be created that meet certain specific requirements or criteria. A specification is delivered using a formal language, in order to reduce or eliminate any ambiguity. An example of a specification is the expression of context-free grammar using Backus-Naur Form (BNF) [202]. In ontology languages, the syntax and the semantics are both given using specifications; this in agreement with Borst [35] who states that an ontology should be formal. Finally
5. *Conceptualization*; referring to the abstraction of phenomenon in the domain of discourse, whereas the ontology formally models the conceptualization [100]. It is

therefore possible to express, logically, the word “reality”. Take, for example, the following propositional logic formula [95]: $\forall A(X) \rightarrow B(X)$, which can be used to express the sentence “all men are mortal” as: $\forall x.man(x) \rightarrow mortal(x)$.

A variety of ontology definitions exist in literature, however, most are derived from Gruber’s definition, meaning his definition is a widely accepted and used definition of ontology. Because of this wide acceptance, the work in this thesis will use the definition of ontology as given by Gruber [97].

3.3 Components of Ontologies

Based on Gruber’s definition of an ontology, as narrated in the preceding section, an ontology is associated with terms or vocabularies that relate with each other (e.g. hierarchical classification) to describe a specific knowledge domain. The fundamental ontology components are: concepts, relations, instances, and axioms [92, 93, 97, 238]. It is not necessary for an ontology to include all of these components - this depends on the complexity, expressiveness and the representation of an ontology [238]. We now describe these components in detail.

Concepts. The concept is considered the backbone of an ontology. It describes the reality of the world, and can be used to describe such things as objects, functions, actions, strategies, or reasoning processes [92]. Usually, concepts are organised as hierarchical taxonomies (Concept, SubConcept etc.) in the ontology [243]. For example the physical activity ontology¹, which is re-used in this research, has ‘Exercise’ as one of the main concepts and sub-concepts like ‘AerobicExercise’, ‘AnaerobicExercise’ and ‘FlexibilityExercise’. Concepts are sometimes referred to as classes [192], particularly for the ontology development process (as in Section 3.6).

Relations. The main goal of the relation is to link concepts, i.e. related terms of the domain. Two main types of relations included in an ontology are object property relations and data property relations. An object property relation is usually used to link two classes together, with one of these classes represented as a domain of the relation and the other as the range of the relation. A data property relation allows assigning of literal values to class instances. Other relations automatically induced by the ontology hierarchical structure are RDF (Resource Description Framework) property relations (Section 3.5.3), such as `subClassOf`, which link hierarchical classes together. Following the

¹<http://biportal.bioontology.org/ontologies/SMASH?p=classes&conceptid=root>

physical activity ontology example introduced in the concepts block above, examples of object property relations include ‘doesRunning’, ‘doesWalk’ and ‘doesWeightLifting’. Examples of data property relations in the same ontology include ‘hasAvgSpeed’, ‘hasIntensity’ and ‘hasTotalStepCounts’.

Instances. Instances, also known as individuals or objects [264], are used to represent individual specific instances of concepts [192]. For example, ‘boxing’, ‘running’, ‘soccer’ and ‘tennis’ are instances of the ‘AthleticSports’ class in the physical activity ontology.

Axioms. Axioms are the assertions [278] (including rules) that are used to represent facts (truthful sentences) in a logical form [88]. Axioms are used to restrict the relations and concepts of an ontology [278]. For example, within the physical activity ontology, ‘AthleticSports’, ‘Exercise’ and ‘OccupationalActivity’ are subclass that are restricted by disjoint axioms to prevent them from overlapping.

3.4 Types of Ontologies

Though there is no agreed taxonomy of ontologies, there are various classification systems for an ontology considered in the literature, which categorise based on either a particular property of the ontology, or a particular property about the domain or task to which it is applied [6]. Mizoughi *et al.* [6, 178] classify ontologies into four types: content, communication, indexing, and meta-ontologies. Then content ontologies are classified as domain, task, and general/common ontologies. Van Heijst *et al.* [256] classify ontologies into two types: type and structure of classification, and the subject of conceptualisation. Guarino [6, 101] classifies ontologies based on their level of dependence on a particular task, distinguishing between top-level, domain, task, and application ontologies. Lassila and McGuinness [6, 141] classify ontologies according to the information need, and richness of the internal structure, distinguishing between the following types of ontologies: controlled vocabularies, glossaries, thesauri, informal is-a hierarchies, formal is-a hierarchies, formal instances, frames, value restrictions, and general constraints. Another ontology’s classification is based on the subject of the conceptualisation [6, 94] to provide four different types of ontologies: domain ontologies, task ontologies, domain-task ontologies, and application ontologies.

We do not seek to present an exhaustive typology of ontologies in this section, we simply look to explain in detail the most common types of ontologies:

Top-Level Ontology. Top-Level ontologies describe very general and high-level concepts, which are related to particular problems or domains. The conceptualizations which

they present are not specific, but can be applied across many individual problems and domains. For example, an ontology presenting space and time [259] can be applied across many problems and domains.

Domain Ontologies. Domain ontologies are commonly used to represent specific domain knowledge (e.g. medical, biology). Domain ontologies use vocabularies and relationships between concepts to express the domain knowledge. Examples include disease ontology² [231] and physical activity ontology³.

Task Ontologies. Task ontologies are used to model domain vocabularies that are relevant to a generic or specific activity (e.g. to solve problems). Task ontologies are not restricted to the single domain – they can be used in modelling different tasks irrespective of the domain. Examples of task ontologies include Scheduling Applications ontology (e.g. Emergency Management Natural Events category with the hierarchy [273]) [217].

Application ontologies. Application ontologies are based on a particular application and describe the vocabularies related to that particular application. An application ontology is extendible, in particular into a specific problem or application. Examples of application ontologies include navigation ontologies [218] (e.g. MEDICO-Annotation-Ontology [233]).

The ontology developed in this thesis is a ‘Task Ontology’, and is built to purposely introduce a classification that has been neglected in current Behavioural models such as Behaviour Change Technique (BCT) [168]. This work proposes a barrier classification ontology which highlights limitations Type 2 Diabetes (T2D) patients suffer during rehabilitation. Barriers to physical activity have been under-researched up until today, which is evidenced by the lack of a barrier classification.

When we consider the different types of ontology, Task, Application and Domain ontologies are all specializations of the Top-level ontology [252]. In a similar manner, Application and Task ontologies are specializations of Domain level ontologies. In this research, in order to create our task ontology, we merge our own systematically defined concepts with concepts derived from four existing domain ontologies i.e. BCT [168], General User Model Ontology (GUMO) [105], physical activity ontology⁴ and disease ontology⁵ [231].

²<http://disease-ontology.org/>

³<http://bioportal.bioontology.org/ontologies/SMASH?p=classes&conceptid=root>

⁴<http://bioportal.bioontology.org/ontologies/SMASH?p=classes&conceptid=root>

⁵<http://disease-ontology.org/>

3.5 Ontology Representation

In order to achieve the ontology’s objectives, an ontology is formally represented in machine-readable languages known as ontology languages. Ontology languages allow users to write explicit, formal conceptualizations of domain models [14] used to represent ontologies. We review three current standards (Web Ontology Language, Description Logic and Resource Description Framework) that are used in developing our proposed ontology below.

3.5.1 Web Ontology Language

Web Ontology Language (OWL)⁶ is a prevalent knowledge representation language with a specification developed and maintained by the World Wide Web Consortium (W3C)⁷. It is a standard to represent an ontology in the domain of the semantic web. OWL allows complex knowledge (expressive sentences), such as “*Every exam must have at least one examiner who is a professor or all examiners of an exam must be professors* ($Exam \sqsubseteq \exists \text{hasExaminer} \cdot Professor$)” to be expressed logically [111].

The ability of OWL to model the complex representation of knowledge, enables it to reason about the given knowledge domain.

Constructs of OWL. Whereas Description Logic (DL) uses terminologies that include concept, property/roles, and individual, OWL uses different terminologies, which are described below.

- *Classes*: a set of individuals with shared properties. Classes start with uppercase letters.
- *Properties*: represents attributes of classes (a data property or concrete role) and the relationships between classes (an object property or abstract role) [83].
- *Individuals*: instances of classes; linked to other individuals using properties.

The OWL family of languages consists of three different sub-languages, which give the user the ability to chose the version most suited to their use case: OWL Lite, OWL DL, and OWL Full [111].

⁶The distorted or illogical acronym of Web Ontology Language (OWL; not WOL), which was suggested by Tim Finin, hides behind it an interesting back story and uncertain origin; it possibly refers to the original character (owl) in the A.A. Milne story, “Winnie-the-Pooh”. Other possible reasons are that owl has no ambiguous pronunciation, that it brings greater opportunity for logo design, and that the animal owl is associated with wisdom [111].

⁷<https://www.w3.org/>

OWL Lite is the most basic of the three variants; it is the least expressive of all variants, decidable, and has exponential computational complexity. *OWL DL* encompasses all features of *OWL Lite*, being more expressive and still decidable. It is supported by the majority of software tools and the computational complexity in the worst-case is worse, at Nexp. *OWL Full* encompasses all features of *OWL DL*, and is the most advanced variant of the language. As a result, it is undecidable, very difficult to work with, with little support from existing software tools.

OWL development is led by the W3C and has had several iterations since its initial launch, with the latest release being OWL 2, released in October 2009 [111]. New constructs introduced in OWL 2 include ones concerned with description of the properties, such as property chain axioms, qualified cardinality restrictions, and negative property assertions [184]. For example, OWL 2 introduces syntactic sugaring for defining multiple disjoint classes in the same axiom (disjoint WITH) [111].

OWL is based on DL and is more formally “equivalent to description logic SHOIN” [266]. There are many examples of OWL applications, especially in public health. One particular public health application of OWL is the human disease ontology⁸ [230], which is further discussed in Section 4.3.3.

3.5.2 Description Logic

Description Logic (DL) is defined as a “family of logic-based knowledge representation formalisms, which can be used to develop ontologies in a formally well-founded way” [18, 245]. This means DL provides “a logical formalism for ontologies” to represent and model the domain concept, and more specifically the semantic web [114, 225]. DLs are decidable subsets of first-order logic, serving as the basis for the Web Ontology Language (OWL) [122]. In Section 3.5.1, we discuss OWL in further detail (including the three sub-languages of OWL).

The basic building blocks of DL are concepts, individuals and roles. The definitions of DL components are identical to those of ontology, as introduced in Section 3.3, though a different naming convention is used: concepts, individuals and roles of DL map to concepts, individuals and properties of an ontology, respectively. The elements of the various representations are capable of representing the reality of the world in a logical structure, called TBox. For example, consider the following two scenarios: a parent is a

⁸<http://disease-ontology.org/>

person (human) that has one or more persons as their child, and a grandparent is a parent that has one or more children, of which one or more is also a parent. DL represents these scenarios as follows [59, 60], (child is also shown to increase understanding).

$$\text{Parent} \equiv \text{Human} \sqcap \exists \text{HasChild} \exists \text{HasChild}.\text{Human}$$

$$\text{Child} \equiv \text{Human} \sqcap \exists \text{HasChild} \exists \text{HasParent}.\text{Parent}$$

$$\text{GrandMother} \equiv \text{Parent} \sqcap \exists \text{HasChild} . (\exists \text{HasChild}.\text{human})$$

The above statements used to represent the scenarios in DL are examples of *axioms*, which are the fundamental modelling concept in DL. An *axiom* is defined as “a logical statement relating roles and/or concepts” [224].

DL is also capable of reasoning using the open world assumption [279], i.e. anything might be true unless it can be proven false [279]. This reasoning plays an important role in designing and maintaining an ontology, as well as answering CQs and queries (Section 5.3).

3.5.3 Resource Description Framework

Despite not formally being defined as a language for expressing ontologies (like OWL), Resource Description Framework (RDF) is a proven standard for establishing interoperability on the web. It enhances the processing and use of meta data [14]. Whilst RDF and Extensible Markup Language (XML) are used alongside one another, they’re conceptually different, RDF is responsible for the semantics whereas XML is responsible for the syntax, as applied in real world computer applications [66]. Conceptually, RDF defines three object types [40]:

1. *Resources*: A resource is anything expressed in RDF, which is an instance of a class. For example, *Exercise* is a class in the physical activity ontology.
2. *Properties*: A property is a specific attribute of a resources. For example, *Prevents* is a property of the *physical activity* and *barrier* resources.
3. *Statements*: A statement is an RDF statement consisting of a specific resource together, a named specific property, and the value of that property for that resource.

Resource Description Framework Schema

Resource Description Framework Schema (RDFS) is an extension of RDF that allows modelling of externally specified semantics, for purposes of mapping them to specified sources. RDFS primitives include `rdfs:Class`, `rdfs:subClassOf`, `rdfs:subPropertyOf`, and `rdfs:label` etc. Each of these primitives has got a unique role in the data model of an RDFS schema, for instance `rdfs:Class` maps to `rdf` expressions; `rdfs:subClassOf` models the

subsumption hierarchy between classes; `rdfs:label` maps to an instance of a property.

Elementary statements are often employed in describing elements of an RDF schema [14]. Using a scenario with the two domains adopted in this work [Physical activity and Barriers], we now give samples of elementary statements. Physical activity and Barrier form classes as `Class(PhysicalActivity)` and `Class(Barrier)`. Sub classes to Physical activity such as *Exercise* and *OccupationalActivity* are organized in class hierarchies as `subClassOf(Exercise, PhysicalActivity)` and `subClassOf(OccupationalActivity, PhysicalActivity)`, respectively. *Walking* and *Running* are *instances* of *Exercise* initiated as `Individual(walking type(AnaerobicExercise))` implying (‘Walking is of type AnaerobicExercise’). ‘Rainy Weather’ is an instance that prevents Walking, this relationship is modelled as `Individual(‘RainyWeather’ value(prevents ‘Walking’))`.

3.6 Ontology Development Process

In order to develop an ontology which allows for sharing and reusing the given domain knowledge, and for making domain assumptions explicit, etc., an ontology development process could be used to guide the development. There are several methodologies [94, 99, 192, 242, 254] which are reviewed in [102] that can be followed to create an ontology from scratch. The difference between these methodologies is often the inclusion and arrangement of particular steps. For example, the methodologies in [99] and [254], stipulate that the objectives of the ontology must be identified before starting. The methodologies in [94] and [192] provide processes to set and limit an ontology’s objectives (e.g. determine the domain and scope of the ontology [192]). Another common difference between methodologies is differing names or terms used to describe steps with similar requirements. For example, the first step, usually concerned with defining objectives or purpose, is described in different ways in different studies (e.g. “determine the domain and scope of the ontology” in [192] and “specification” in [120]), although the step’s requirements are the same. Furthermore, some methodologies focus on creating an ontology in a specific field, such as a methodology to create a chemical ontology [146].

In the ontology development process, it is advisable to follow one of these methodologies in order to ensure that all steps are achieved to meet requirements and goals of the ontology, especially if the objectives of the ontology are not clear from the start [41]. Adapting any one of the methodologies that makes it relatively easy for both domain and non-domain experts as well as researchers to review, re-use and understand the ontology. Similar to

other researchers, in addition to utilizing these methodologies we ensure to do the following:

- (1) Endeavour to construct the ontology such that it reflects the reality of the world and
- (2) Iteratively follow the development steps i.e. iterative ontology designing which is crucial during the entire life cycle of the ontology.

When establishing an ontology, it is possible to merge different methodologies to create a new ontology, where each methodology provides a distinct advantage towards achieving the design ideas [41].

In this section, we present details about one of the most common methodologies used in ontology development, called “Ontology Development 101” [192], also referred to as the “101 Methodology”. This methodology has been applied practically in different ontology applications, particularly in health. At the time of writing, the number of citations of this methodology has surpassed 6400. This gives a positive indicator of its success, wide adoption and application in various studies and research areas. The 101 methodology includes quick and simple steps that can guide in creating an entire ontology from scratch, even if the objectives of the ontology are ambiguous, vague or unclear at the beginning. The 101 method is easy to understand and apply, even for non-domain experts [44]. We discuss the 101 methodology below, which is later applied practically in Section 4.3 to develop the Barrier Ontology.

1. **Decide the domain and scope of the ontology.** Determining the ontology’s domain and scope is commonly the first step in the ontology development process. The answers of several fundamental questions are required to define the domain and scope of the ontology, though the response to these questions could change during the design process. These questions include, but are not limited to, the following:

- What is the area covered by the ontology?
- What is the ontology used for, and to whom is the ontology targeted (end users)?
- To what types of questions should the ontology provide answers?

Noy *et al.* highlighted the importance of competency questions (CQs) (at a later stage of the ontology development or preferably upon completion) indicating that they serve as the litmus test for the ontology [192]. In other words, they’re essential in evaluating how well the ontology fits the decided domain and scope. The CQs are defined as “the definition of ontology requirements described as informal questions that an ontology must be able to answer” [56]. This means that an ontology that is fit for purpose should be capable of answering these CQs. Some of the CQs that are related to this study (barriers to physical activity behaviour) are shown in Chapter 5

(Section 5.3.2).

2. **Import existing ontologies as much as possible.** Usually, the second step of an ontology development process is to import the current existing ontology. It is essential to check whether there is an existing ontology in the domain area that can be imported, instead of creating one from scratch. It not only saves effort on building an ontology from the beginning, but is also good practice in the ontology development process. This step is usually necessary if the given ontology is required to link with other related ontologies, such as when using existing control vocabularies. Many different existing ontologies or enumerate terms are either free to use or open-sourced, available via online medical resources such as Medical Subject Headings (MeSH) [197] and the National Center for Biomedical Ontology (NCBO), which has a portal called BioPortal⁹. These existing ontologies are often presented in the RDF format. Sourcing suitable and reliable existing ontologies is a significant challenge for ontology developers. Once a suitable existing ontology is found, unless it's regularly updated, it will often contain either out-dated or different formalisms that are inconsistent with current and future standards in web semantic development. Nonetheless, the process of transforming an ontology from one form to another form is not a complicated process [192].
3. **Enumerate important terms in the ontology.** This step can otherwise be thought of as the vocabulary collection stage, as it's during this stage that terms or vocabularies are gathered from different sources. These sources could include domain experts, interviews with intended users, questionnaires, and literature reviews. During this process, it is preferable to identify as many terms as possible, regardless of any overlap among terms, relations and properties. Later, some of these terms become relationships (properties) or individuals.
4. **Define the classes and the class hierarchy.** The fourth step of the ontology development process defines the classes and hierarchy or taxonomy relationships between related classes (terms). This step selects terms that can be classes, ignoring others such as property terms, which can be handled in a separate step.
In order to develop the taxonomy relationships among the classes, there are several methods that can be followed.

- **Top-down method:** This method starts with the most general concepts in the

⁹<https://www.bioontology.org/>

given domain, which are then broken down into more specific concepts. This is then recursively repeated until a particular granularity is achieved. For example, the human disease ontology begins with the *disease* class (general category). It then proceeds with classifying the *disease* class into eight types of disease, such as *metabolism*. Next, the *metabolism* class is classified into *acquired metabolic disease* and *inherited metabolic disorder disease*, and so on, to reach a specific definition type of disease, such as *diabetic ketoacidosis* (this is demonstrated in Figure 3.1). The top-down method is adopted in this thesis i.e. we initially determine top level concepts which include various barrier types, such as environmental barrier and personal barrier. Environmental barrier is subsequently broken down to weather condition, which is further broken down into the most specific concepts, such as cold, hot, and rainy.

- **Bottom-up method:** The bottom-up method is the reverse of the top-down method. It starts from the specific concept and goes up until the general concept is reached. In the disease model, for example, the classification starts from *diabetic ketoacidosis* until the disease concept (general concept) is reached (Figure 3.1).
- **Combination method:** This model combines the top-down method and the bottom-up method. Ideally, one may optionally begin with the more general concepts, especially the most prominent or salient concepts, thereby obtaining top-level concepts. This can subsequently be followed with identification of mid-level concepts that are relatively less prominent than top-level concepts, finally identifying the least prominent or otherwise most specific concepts at the bottom-level. The ontology hierarchy gradually takes shape with identification of concepts that lie in and amongst the three levels which can randomly follow a top-down or bottom-up method. For example, in building the disease ontology (demonstrated in Figure 3.1), disease of mental health and disease of infectious agent could be identified first as top-level concepts, then amyloidosis and Type 2 Diabetes (T2D) are identified as mid-level concepts; finally, gestational diabetes and prediabetes syndrome identified as the most specific concepts. All the other concepts are then added to the ontology hierarchical structure following either a top-down or bottom-up procedure.

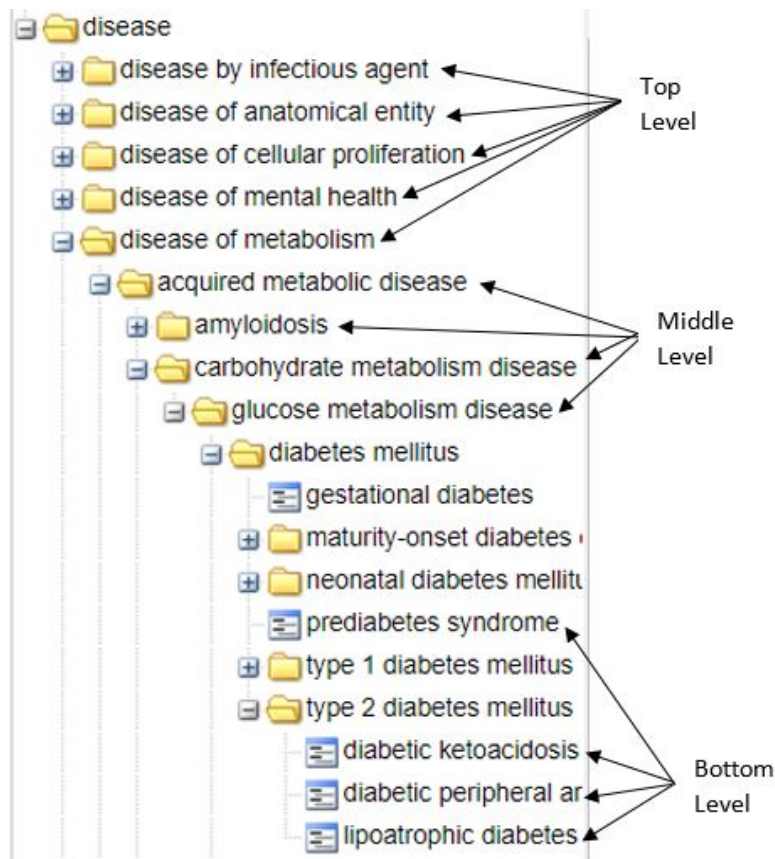


Figure 3.1: An excerpt showing concepts hierarchically organized across the different levels¹⁰.

5. **Define the properties of classes.** In the ontology development process, there are two different types of properties: data properties and object property. This fifth step of the ontology development process considers the data properties of the classes that were defined in the previous step; object properties are discussed in the next step (define the restriction of the classes). Therefore, the property term in this step refers specifically to a data property. These properties are used to define the classes, and to provide necessary and sufficient conditions for class membership [24, 220]. Usually, the terms remaining from the define classes step (define the classes and the class hierarchy) are regarded as the data properties, and are to be classified within their

¹⁰<http://disease-ontology.org/>

particular classes. For example, *name*, *ID*, and *address* are properties that can be used to describe the student class.

6. **Define the restriction of the classes.** There are two types of restriction that are identified during this step. The first is the value type of the class, such as *string*, *number* and *boolean*. The second is the relations between classes, which would be either binary relations or semantic relations. The semantic relation is defined as a “meaningful associations between two or more concepts, entities or sets of entities” [130], such as objects and instances.
7. **Created instances.** The last step of the ontology development creates instances of classes. This requires the following procedure: (1) choosing a class; (2) creating an instance of the selected class; (3) filling the property values of this instance.

It should be noted that the process to create an ontology is not fixed. This means that the order of the ontology development process is different for each situation; the order that is followed above is the most appropriate for our case.

In order to develop an ontology, an ontology editor tool can be used. Protégé is a tool that “allows users to visually manipulate, inspect, browse and code ontologies and support in this way the ontology development and maintenance task” [123, 253], and is one of the most widely used development platforms for ontology-based systems [134]. In this work, the Protégé editor is used to develop the Barrier Ontology (Section 4.4). In addition to Protégé, additional tools exist such as Pellet, RacerPro, FaCT++ and HermiT [111], which support the OWL (Section 3.5.1) and RDF (Section 3.5.3) web semantic standards.

3.7 Ontology Evaluation

The ontology evaluation process serves developers, where it guides them to improve the structure and the results of the ontology. An ontology’s evaluation is one of the most controversial topics in the field of ontology. Many perspectives, techniques and methods of ontology evaluation are presented in the literature; each view is suited to different ontologies, depending on the type and purpose of the ontology. In this work, two different evaluations approaches or techniques are applied to evaluate the Barrier Ontology: the data-driven approach, and the competency questions (CQs) technique.

This section discusses different approaches to evaluate an ontology, focusing on the data-driven approach and CQs technique, which are used to evaluate the Barrier Ontology.

Data-driven. Brewster *et al.* [38] proposes a data-driven (corpus-based) approach,

which compares the given ontology with a data source (corpus) from the knowledge domain of interest, in order to measure the degree of structural fit between them. Further details about this approach and its use in the evaluation of an ontology are presented in Chapter 5 (Section 5.2) before using it to evaluate the Barrier Ontology.

Competency Questions (CQs). The CQs play an important role in both determining the domain and scope of an ontology during the ontology’s creation, and in evaluating the completeness and consistency of an ontology. Since ontology evaluation is the task of assessing the ability of an ontology’s content to achieve its task or requirements, the CQs can be applied to evaluate the ontology. The transformation of the CQs from natural language to formal machine language (e.g. queries) is usually a manual process [274]. During the evaluation process, the returned answers of the queries (which consist of concepts or instances of ontology) must correctly match the natural language answer of the original CQs [224].

Besides the two aforementioned approaches, there are other approaches that can be used to evaluate an ontology. These approaches, which are briefly detailed below, are ‘Golden Standard’, ‘Application-Based’, ‘Assessment by Human / Domain Experts’ and ‘Model Publishing’.

Golden standard. Maedche and Staab [151] suggests an approach, known as “ontology alignment or ontology mapping”, whereby the ontology to be evaluated is compared against a “golden standard”, which may itself be an existing relevant ontology that has been created using reliable knowledge and domain experts.

Application-Based Evaluation. Porzel and Malaka [208] give an application-based approach to ontology evaluation, whereby the ontology is plugged into an application, evaluating the resulting output.

Assessment by Humans or Domain Experts. Lozanotello and Gomez-Perez [148] provide an approach where humans assess how well the ontology performs against a set of predefined criteria, such as prescribed standards or legal requirements, etc.).

Ontology Publishing. Ontology publishing requires that a model of an ontology is published in a formal language (e.g. OWL). This method allows validation [250] and peer review of the ontology. One of the disadvantages of this technique is that it requires a long time to execute completely. For this reason, this type of evaluation has been avoided during this work, but it may be carried out in the future. More detail about this type of evaluation is presented in Chapter 6 (Section 6.5) during discussion of future work.

3.8 Ontologies in Behavioural Health

Before moving to the description of the process that we used to build the ontology in this theses, we conclude the review chapter on ontologies with an analysis of existing ontologies on behavioural change.

We have already discussed how the taxonomy on BCTs, as an example of a recent cooperative effort to create observable and replicable interventions to influence behaviour and health, with their hierarchical definitions and examples, is a step towards an ontology [165], to support the aggregation of behaviour and behaviour change knowledge as well as sharing and reusing useful sources of behaviour knowledge. We also observed that to our knowledge there currently exists no publicly available conceptual classification that models barrier to behaviour change. Having said that, several classifications (taxonomies and ontologies) related to human behaviour exist in available literature. We now review some of these classifications.

3.8.1 Behaviour Change Techniques as Hierarchical Taxonomy

Now that we have explained what ontologies are and how they are built, we can now go back to the Behaviour Change Techniques taxonomy (BCT) [168, 171] and discuss it from an ontology perspective.

The BCT was published in 2013 by Michie *et al.*, [168] and has not yet been updated. As explained in Chapter 2, it includes 93 techniques that are classified into 16 categories, such as feedback and motivation [189]. The BCT is still in its early stages, and currently includes a defined domain and scope, controlled vocabularies or terms, and relations among the concepts or vocabularies. The authors foresee a progress in BCT towards building a behaviour change ontology, “*specifying relations between BCTs, mechanisms of action, modes of delivery, populations, settings and types of behaviour*” [165, 189]. The components of BCT (mechanisms of action, delivery, populations, settings and others) are discussed in detail in Section 2.3.1.

The BCT taxonomy is therefore the most prominent and notable work towards the aggregating and sharing of the behaviour domain knowledge base, hence towards the realisation of an ontology of concepts related to behaviour change and maintenance. It encompasses physical activity behaviour, diet behaviours, smoking cessation, and changing professional behaviour. The BCT can be (and is anticipated to be) extended and updated based on international, interdisciplinary consensus. The BCT is reusable and extendable to

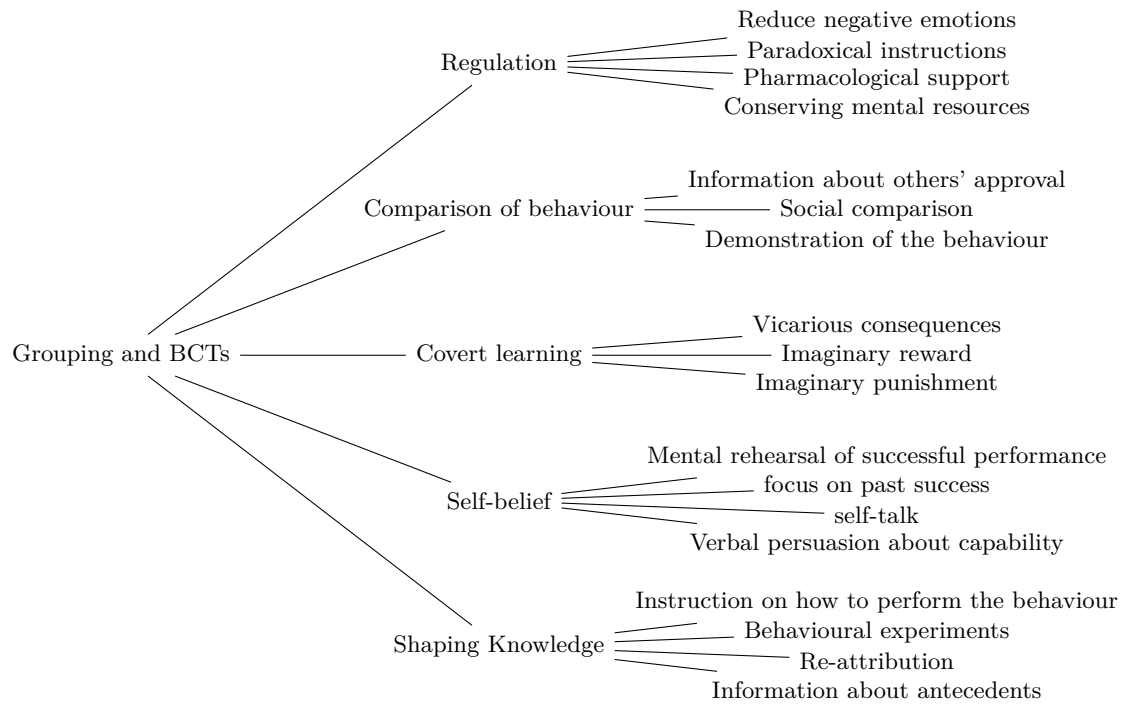


Figure 3.2: A partial view of BCT Taxonomies.

support or to link other hierarchically structured taxonomies or controlled vocabularies. Examples of the techniques included in the taxonomy are ‘review behaviour goal’, ‘goal setting (outcome)’ and ‘demonstration of the behaviour’. Figure 3.2 shows examples of interventions in the BCT taxonomy (partial view) [168, 171].

The BCT is not presented as an independent component, but is included within a machine-processable model (ontology) for behaviour change interventions (Figure 2.4), as mentioned during the discussion of intervention components. As noted above, the BCT, which refers to the intervention in the ontology model, is associated with the behaviour component, that is the targeted behaviour in the intervention’s characteristics, via the mechanisms of the action component. Each component of this ontology model is presented in its own ontology or taxonomy. This causal model presents a structure mechanism, providing a better understanding about behaviours and the behaviour change domain, for example how the mechanism of relations among components work and associate. The components and relationships enable a strategy to extend to other related ontologies or taxonomies. This model with clear definitions, labels, and examples of BCT, is therefore

an important starting point to those interested in developing an ontology in this domain.

3.8.2 Human Behaviour Taxonomy

Another important effort to mention is the human behaviour taxonomy from the World Health Organisation (WHO) [140, 199]. These taxonomies of human behaviour have been developed based on the knowledge of the WHO and on the International Classification of Functioning (ICF), Disability, and Health. These taxonomies include a full definition of their classes, based heavily on the U.S. National Cancer Institute (NCI) Thesaurus, as well as the Oxford English Dictionary [268]. Figure (3.3) shows a part of the behaviour ontology's taxonomy [140], though due to the complexity of the subject matter there is currently no complete ontology capturing all aspects of human behaviour.

3.8.3 Semantic Mining of Activity, Social and Health Data

Another computational effort in capturing human behaviour is the Semantic Mining of Activity, Social and Health Data (SMASH) [206]. This is a deep-learning project which looks to provide predictions on human behaviour, as well as explanations for the predictions, and is based on an Ontology Restricted Boltzmann Machine [265]: a bottom-up algorithm learns user representation from health ontologies, using the user representation to incorporate self-motivation, social influences, and environmental events into the generation of prediction (and explanation of prediction) of human behaviour. The predictions, which model human behaviour, are dependent on social networks, such as self-motivation, social influences, and environmental events. Self-motivation is learned or captured from history and from existing characteristics of the user. The social influences include users' friends and social networks. The environment events are identified through users' social relationships, unacquainted users and social contexts [49]. The explanations of the predictions are useful to increase the reliability of the intervention in changing a specific behaviour. For example, physical activity behaviours such as walking and running are included into SMASH as a type of intervention, with measurements from a device taken every 15 minutes, in order to report the number of walking and running steps.

3.8.4 The Neurobehavior Ontology

The Neurobehavior Ontology (NBO) [89] is an ontology on the domain of Behavioural processes and phenotypes, which are related to behaviour and behaviour change [227].

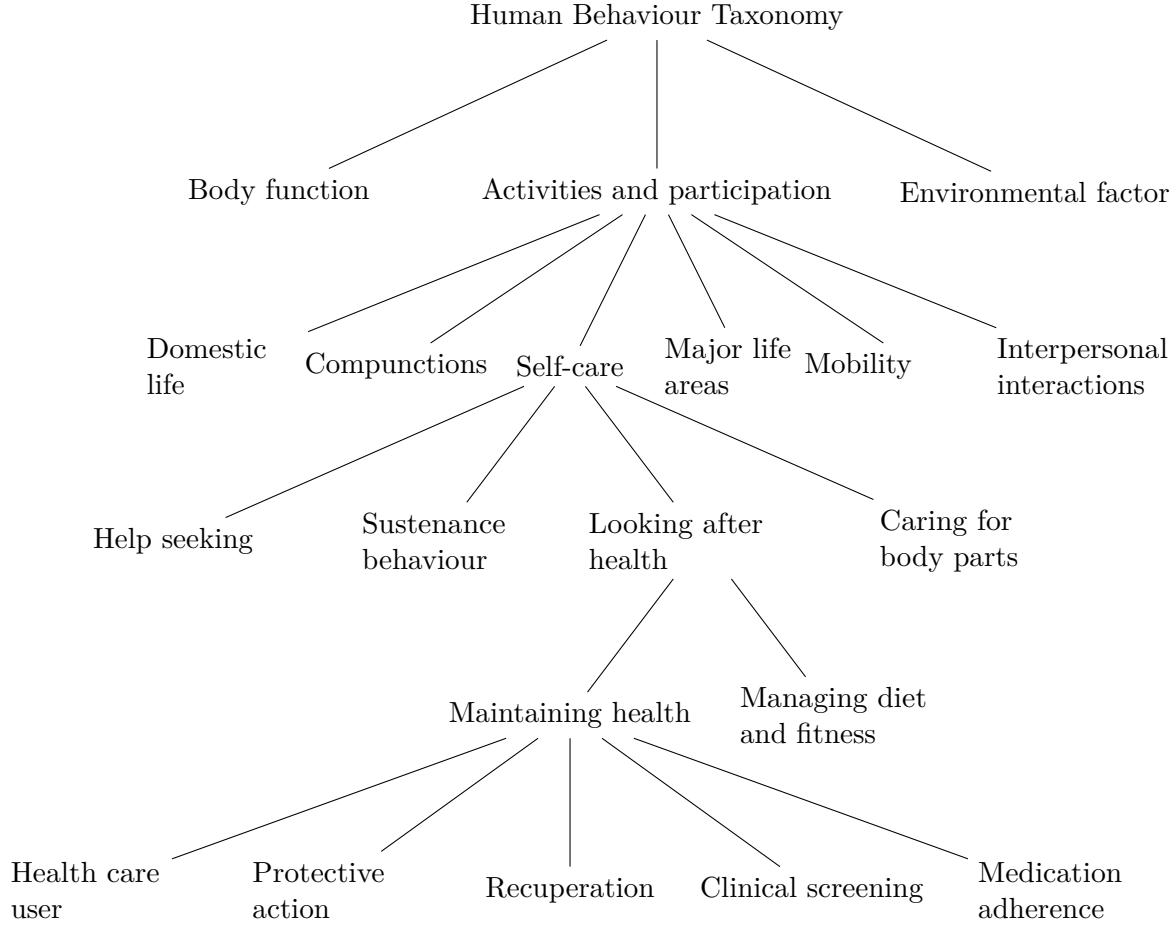


Figure 3.3: Partial view from Human Behaviour Taxonomy [140, 199].

The NBO includes two main components: (1) behavioural processes ontology, and (2) behavioural-phenotypes ontology. The first component of the NBO, the largest, classifies the behaviour process to complement and extend the Gene Ontology (GO), “a major bioinformatics initiative to unify the representation of gene and gene product attributes across all species” [54]. The second component of the NBO classifies normal and abnormal behaviour of organisms. The NBO is extended to include not only the GO, but also other ontologies, such as Uberon, an ontology of anatomical structures in animals¹¹, by using the relation of *by-means*. The ‘motivation behaviour’ and ‘social behaviour’ in NBO classes

¹¹<http://uberon.github.io/>

further classify into *avoidance behaviour*, *thirst motivation behaviour* and *thermoregulation*, etc., *agonistic behaviour*, *communication behaviour* and *group behaviour*, etc., all of which are examples of behavioural process classes [89]. Applications of NBO cover both human behaviour and that of some animals, and it includes 671 classes, 19 individual and 61 priorities. It is freely available for download and use from the BioPortal repository, and regarded as reliable to use [189].

3.8.5 Health Behaviour Change Ontology

The Health Behaviour Change Ontology (HBCO) is one of the closest ontologies to our efforts, and it was built for a project aiming to establish an automated dialogue between a psychologist and a user to provide behavioural counselling [32]. The HBCO ontology has strengthened the linkage between theoretical and practical parts, but few practical implementations exist [241], meaning there is currently no solution to provide a reusable behaviour change ontology. Developed for use in delivering an automated counselling session, this ontology is based on social cognitive theory [20] and the Transtheoretical Model (TTM) we describe in Section 2.2.2, yet we note that the relations between the stages of change in TTM are not indicated in the HBCO. The HBCO therefore models the mental state and emotion of the user, by integration with related behaviour ontologies [128]. The knowledge base of HBCO is acquired from the concept's review of behavioural medicine, and from human experts in areas such as psychology, behavioural medicine, and computerised interventions. The ontology is created in the Protégé editor using OWL. The overall structure of health behaviour in this ontology uses three dialogue-level structures: counselling dialogue, motivational interviewing dialogue, and social cognitive counselling dialogue. Each of these dialogue-level level structures is then classified further into sub-structures. For example, counselling dialogue is classified into social rituals, review tasks, assign tasks, and so on. This ontology can be reused to provide behaviour change advice, to prompt a user to engage in a particular activity. For example, the user may choose their preferred type of exercise (e.g. walking) and then enhance the exercise by means of an intervention, such as signing up for a walking club, or thinking about the value and benefit of regular walks.

3.8.6 User Modelling Ontologies

As our conceptualisation must refer to the model of the person for whom the behavioural change techniques need to apply, it is useful to explore the ontologies which explicitly models

a "user", in terms of their profile, characteristics, and sometimes also their behaviour. One example of a user ontology is the General User Model Ontology (GUMO) [105], which we will also import in our ontology, and which therefore will be further discussed in Section 4.3.3. Beside the GUMO, there exists several other ontologies that encapsulate wider aspects of user (human) activities. One of these, which is in fact a GUMO extension, is the User Navigation Ontology (UNO) [132], which constrains user-modelling to just a single activity which is Navigation and wayfinding. Another user profile ontology is OntoPIM (Ontology Personal Information Management), which describes various users' dimensions and shares a lot of concepts with GUMO [153]. OntoPIM describes the user's domain of interest covering user characteristics such as personal information, general user characteristics, user abilities, preferences (i.e. abstract likes/dislikes), interests, activities and profession, which are rather static information. It is to be noted that GUMO remains the ontology of choice for many reasons, mainly because of its fluidity when exchanging user model data across different user-adaptive systems [106], and because it adopts a global semantic web language OWL [63].

3.8.7 Ontologies of Physical Activity Behaviour and Exercises

Finally, in order to address the specific domain of choice, that is physical activity behaviour, we need to mention some behaviour ontologies related to the physical activity or exercise domain, in particular the Exercise Ontology [209] and the Ontology of Physical Exercises¹². The former was built to support a recommendation system that suggests physical activities to diabetic patients using data such as weight and height. This ontology however only models exercise and neglects to model more general physical activities, such as daily house chores. In addition, the authors of this ontology did not provide a public resource for downloading the ontology. The ontology of physical exercises [209] on the other hand, maps exercises to their expected health outcomes [72]. Another example of an ontology which comprehensively covers physical activities is the Physical Activity ontology¹³, which captures exercise and non-exercise activities. We re-use the physical activity ontology in our work to model physical activity concept, as discussed in Section 4.3.3.

None of the aforementioned works however cater for the actual challenges faced by users who have endeavoured to employ the intrinsic behaviour strategies they entail. That

¹²<http://bioportal.bioontology.org/ontologies/OPE/?p=summary>

¹³<http://bioportal.bioontology.org/ontologies/SMASH?p=classes&conceptid=root>

is to say that the works that we have reviewed all assume that the user is completely compliant with the advice given and that they face no issue in carrying out the tasks advised. Management of these challenges and barriers is crucial for users attempting to change, rehabilitate or maintain their behaviour. It is this challenge that motivates and inspires the work covered in this thesis. The main objective is to develop a classification (ontology) that would enhance real-world employment of behavioural techniques stipulated in such classifications. In addition, we also look to enrich the knowledge base of human behavioural challenges, which is potentially useful for other researchers and health behaviour therapists.

3.9 Summary

The emphasis in this chapter is to exhaustively discuss what an ontology is, the components it entails, and different techniques of evaluating an ontology. The chapter began with definitions of an ontology as perceived by the both the Philosophy and computer science domains. Key ontology components are broken down describing their significance in ontology modelling, these include concepts, relations and instances, followed by a discussion on the types and different ways of representing an ontology. A detailed description of the widely adopted "Ontology Development 101" methodology is included in this chapter. Finally, the chapter reviews many existing ontologies with a keen focus on behaviour change and behaviour change maintenance. The contents of this chapter form therefore the basis for not only understanding the research questions related to this thesis, but also for understanding the motivations behind the structure of the Barrier Ontology.

We are now ready to move to the description of the methods used in constructing the Barrier Ontology (Chapter 4), main contribution of this thesis.

Chapter 4

The Barrier Ontology Method and Development Process

4.1 Introduction

Chapter 3 discussed the areas of ontology (definition, components, types, development process and evaluation) that are relevant to building the Barrier Ontology. Chapter 3 also reviewed various existing ontologies pertaining to behaviour.

This chapter illustrates the development process of the Barrier Ontology. The chapter starts by discussing both functional and non-functional requirements of the ontology in the context of T2D, before presenting the conceptual barrier model (Section 4.3), and then concluding with an extended sequential development process (Section 4.4) of the Barrier Ontology. The chapter addresses the main thesis research question: *“How can we use ontologies to formalise the notion of barriers to behaviour change and their underlying assumptions in a machine-readable format to support health informatics applications?”*, which is framed by four sub-questions: (a) *“How can we build an ontology of “barriers to physical activity for T2D patients?”*; (b) *“How can we demonstrate the use of a formal methodology, including the notion of ontology reuse, to objectively support the Barrier Ontology?”*; (c) *“How can we use the Barrier Ontology to produce suitable recommendations of physical activities which take into account barriers to such physical activity from T2D patients?”*; (d) *“How can we evaluate our methodology, and offer general suggestions for ontology developers?”*.

We answer sub-question (a) in two stages; stage 1 (domain and scoping, systematic review

and barrier concept) (Section 4.3) justifies the selection of a specific behaviour (physical activity) and demographic group (T2D patients) as well as determine the terms related to these domains; stage 2, which is the development process of the Barrier Ontology (the steps which decide the domain and scope of the Barrier Ontology, which considers reusing existing ontologies, and which define the classes, the class hierarchy and the properties describing these classes) (Section 4.4) constructs the Barrier Ontology using concepts and relations gathered from stage 1.

In order to answer sub-question (b), we conduct an analysis of existing ontologies in order to determine the reuse of existing models (physical activity, human disease, user and Behaviour Change Techniques) in Section 4.3.3, and selecting the existing concepts to reuse (e.g. user or patient, physical activity and disease). We develop our ontology following the 101 ontology development process (Section 4.4), which incorporates the re-use of existing ontologies.

In order to answer sub-question (c), we define two main relations during the scope-definition (Section 4.3.1) and term-collection (Section 4.3.2) stages of the ontology, these are *prevents* and *isSuggestedFor* (these are inverse of each other) (Section 4.3.4). Both relate the physical activity concept and the barrier concept, however the latter inherently depicts that a given physical activity is suitable in limiting a given barrier, thereby answering sub-question (c).

The answer to sub-question (d) will be discussed in a dedicated Chapter 5.

4.2 Requirements

According to [244], the ontology requirements specification should state why the ontology is being built, what its intended uses are, who the end-users are, and which requirements the ontology should fulfil. With this in mind, we can say the following about the requirements for the *Barrier Ontology*: The Barrier Ontology aims to create a conceptual barrier model (Barrier Ontology) to support health behaviour applications, such as Behaviour Change Technique (BCT). There is currently no barrier ontology or hierarchical taxonomy to support behavioural knowledge extraction, despite barriers being mentioned by different studies (Section 2.5).

In order to contextualize, investigate, and define the problem, a particular scenario was selected, i.e. barriers to physical activity behaviour. Regular physical activity not only improves the general health of the general public, but also aids in the effective management

of some common diseases, such as T2D. We therefore looked to identify a chronic disease that is both prevalent worldwide [15] and can also be effectively managed by regular physical activity. So our focus is on physical activity behaviour for patients with T2D. T2D has a significant impact on the patient's lifestyle, but can be managed, among other things, by a healthy physical activity regime.

The importance of regular physical activity is perhaps not what people naturally associate with effective management of T2D (where insulin and diet immediately come to mind), hence it is likely to be a potential source of barriers. This is discussed in more detail in Section 4.3.1. The ontology aims to model the knowledge used by a decision support system whose intended end-users are patients with T2D, their healthcare providers (e.g. doctor, nurse, or psychologist), and interested researchers in the fields of health and behaviour and that exploits the relationships existing between barriers, activities and patients/users

The four main functional requirements of the Barrier Ontology are: (1) to identify the barriers, based on user's characteristic (e.g. personal information, such as employment status), including circumstances that may prevent the user from performance of physical activity (e.g. environmental condition such as weather and facilities, and social conditions such as social relations and support); (2) to recognise barriers that prevent performance of a specific type of activity; (3) to suggest physical activities to limit identified barriers; (4) to retrieve the list of smaller specific barriers that inherently belong to the six main categories.

We formulate four templates of competency questions (CQs) (Section 3.6), to fulfil the four requirements of the Barrier Ontology. Each template contains four different competency questions, and each of these inherently satisfies the template's corresponding functional requirement. Table 4.1 contains a sample of the mapping of competency questions to functional requirements. Additionally, Chapter 5 contains an expanded version of this table (Table 5.2); this includes other components that were relevant in evaluation of the ontology, such as questions responded to by domain experts in an online survey used in the evaluation described in Chapter 5.

Table 4.1: A sample of the mapping of competency questions to functional requirements.

Functional requirement	To recognise barriers that <i>prevent</i> performance of a specific type of activity.
------------------------	---

Continued on next page

Table 4.1 – *Continued from previous page*

Template 3	<p>CQ 3.1: What are the barriers that prevent or limit people from performing ‘<i>Football</i>’, despite being interested in this activity?</p> <p>CQ 3.2: What are the barriers that prevent or limit people from performing ‘<i>Swimming</i>’, despite being interested in this activity?</p>
Functional requirement	To suggest activities to limit identified barriers.
Template 4	<p>CQ 4.1: What physical activity is suggested for the bad weather condition (‘<i>Raining</i>’) barrier?</p> <p>CQ 4.2: What physical activity is suggested for the ‘<i>Financial problem</i>’ barrier?</p>

Other non-functional requirements that we seek to fulfil in the Barrier Ontology include building an ontology that is reusable and extendable [244]. (Section 5.3.2) by running different queries (CQs) and retrieving results which contained no consistency among barrier’ classes. We therefore want to ensure that the Barrier Ontology is suitable for reuse in different domains. This would allow the future reuse or importing of the Barrier Ontology and for the encompassing of more terms, concepts, and relations, or to include additional behaviours, such as nutrition behaviour.

4.3 Conceptual Barrier Model

We have already discussed the BCT taxonomy as an ongoing collaborative efforts to create a behaviour ontology to support behaviour change, but we note here again the main concepts as we move to the realisation of the ontology. Different barriers prevent Behaviour Change Techniques from achieving successful behaviour change and maintenance [171]. These barriers can be defined as “perceptions concerning the unavailability, inconvenience, expense, difficulty, or time consuming nature of a particular action” [157, 204]. Examples of these barriers are personal barriers [75], health barriers, environmental barriers, and psychological barriers [198]. A conceptual model of barrier and behaviour must effectively support software applications in selecting effective interventions to influence.

A high level such model is depicted in Figure 4.1: this model reflects the rationale of the earlier introduced requirements, while simultaneously incorporating prior knowledge of behaviour and behaviour change, as discussed in Chapter 2. In the following sections we

will detail the process we have followed to arrive to this model.

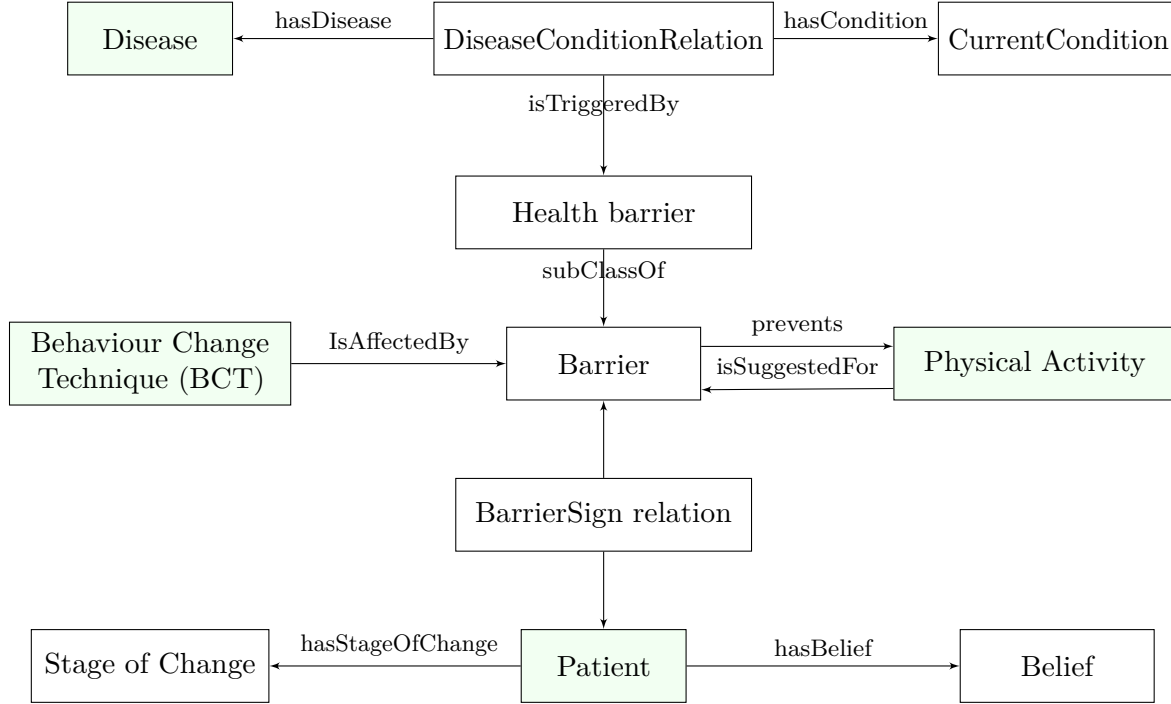


Figure 4.1: A high-level conceptual model of barriers to physical activity behaviour for T2D patients.

4.3.1 Domain and Scoping

As discussed in Section 4.2, we have selected physical activity behaviour for T2D patient as a specific scenario to focus our discussion.

T2D is the most common type of diabetes, accounting for about 95% of cases [16]. T2D occurs when the pancreas cannot function, or cannot produce enough insulin, which helps the body's organs to absorb glucose from the blood. The number of people living with T2D has increased dramatically from 108 million in 1980 to 422 million in 2014 [200]. This number is expected to reach 552 million in 2030 [235] and 592 million in 2035 [109].

More than two million deaths occur each year as a result of diabetes and its associated complications [200], such as nerve damage (neuropathy) and kidney damage (nephropathy). Annually, diabetes management costs about 10% of the entire health budget. This is estimated to reach 17% of the health budget in 2035 as reported by [109].

There are different types of diabetes which include, type 1 diabetes, T2D, gestational diabetes, mellitus and monogenic diabetes syndromes [16, 200]. These different types are less common, so are not discussed as they are outside of the research domain.

T2D, also known as ‘non-insulin-dependent diabetes’, causes any insulin present to be used inefficiently or causes the body’s cells to inadequately absorb insulin [16, 200]. This means that there is insulin present, but it is ineffective. The physical activity supports in activating this insulin, and absorption of it, by the cells of the body. It follows therefore that an inactive diabetic patient (who partakes in little or no regular physical activity) is more prone to diabetic complications in the future. Beside regular physical activity, several other vitally important healthy behaviours include healthy nutrition and smoking cessation, which supplement medicating the condition with insulin injections or tablets.

Physical activity defined as “any bodily movement produced by skeletal muscles that lead to energy expenditure” [45]. Consequently, the physical activity not only includes the exercise behaviour type but also includes normal daily activity behaviours, such as home activities (e.g. cleaning and sweeping, climbing up stairs, walking the dog, etc.). Examples of physical activity include athletic sports (e.g. soccer, basketball, and boxing), aerobic exercise (e.g. running and walking), and occupational activity (e.g. lawn and garden, plumbing, and masonry) [2].

Different barriers limit engagement or participation in physical activities, more so regular physical activity [22, 25, 75], and we attempted to capture them through a systematic review exercise described in what follows.

4.3.2 Systematic Review

To our knowledge, there is an unavailability of models conceptualizing our pre-determined concepts, particularly the barrier concept. This is affirmed through a scoping review (of material that dates back to 1999) of prolific online research repositories such as Google Scholar, ACM Digital Library, PubMed and IEEE-xplore, as well as ontology repositories like BioPortal¹, and the Open Biomedical Ontology (OBO) Foundry². It was plausible that prior literature might have some information about barriers to physical activities necessary for this work. We therefore conducted an extensive systematic review to not only acquire barriers’ vocabularies or terms, but also to search and find physical activities suitable in limiting these barriers. During the systematic review, we embarked on appropriately classifying

¹<https://bioportal.bioontology.org/ontologies/D0ID>

²<http://www.obofoundry.org/ontology/doid.html>

the terms to form a hierarchical structure that would drive the ontology construction (Section 4.3.3).

Method. Six electronic databases were used to identify relevant studies across the behavioural science and health fields: IEEE-Xplore, PubMed Database, Google Scholar (first 15 pages of search results), PsycInfo database, and ACM Digital Library. The search was limited to fetch papers and articles published after 1 January 1999. We believe that this is a sufficient time-frame on which to carry out the systematic scoping review of barrier terms or vocabularies.

Combinations of the following keywords were used to query the repositories: ‘barrier’ or ‘obstacle’, ‘physical activity’ or ‘exercise’, and ‘T2D’ or ‘diabetes’. The results of the queries contained relevant studies for further review. We now give two examples that show how we combine these keywords to identify and retrieve relevant articles. The first example searches with the keywords ‘Barrier’, ‘physical activity’ and ‘T2D’ in the search engine of ACM Digital Library; the result contains 59 retrieved studies. The second example swaps ‘barrier’, which is the most common term used in the literature, with ‘obstacle’, in order to capture studies that use ‘obstacle’ instead of ‘barrier’. This query returned 13 studies. After removing the duplicates or studies that co-exist across the results from these different search methods, 62 studies are retained. Figure 4.2 shows a flow chart of the study retrieval process obtaining the final result after removing duplicates from the two queries. The other electronic-databases are queried using the same process detailed in Figure 4.2.

Criteria for excluding studies. We exclude studies that meet one or more of the following criteria: (1) barriers to another health behaviour (e.g. nutrition); (2) barriers to physical activity for another disease (such as type 1 diabetes); (3) publications in a language other than English; (4) publications dated before 1999; (5) Early access articles or unpublished studies; and partial studies.

Finding. From 862 identified studies whose title and abstract were each reviewed, we remove duplicates and retain 794 studies for reviewing and screening. 753 studies are excluded because they do not match the selected domain and range of our study. For example, studies narrating barriers for different health behaviours (e.g. nutrition), and studies narrating barriers to physical activity for different diseases (such as cancer). 41 studies are then eligible for screening. During screening, we search for studies that jointly discuss T2D and physical activities. Therefore, 22 studies are rejected: 17 studies [17, 52, 103, 133,

144, 150, 174, 182, 185, 207, 229, 236, 246, 248, 249, 255, 257] do not focus on barriers to physical activity and 5 studies [46, 73, 183, 201, 232] do not focus on T2D. This results in 19 studies [12, 22, 25, 29, 34, 70, 74, 75, 108, 136, 137, 142, 143, 157, 159, 172, 203, 237, 240] that meet the inclusion criteria (inclusion within the scoping review). These criteria are demonstrated in Figure 4.2.

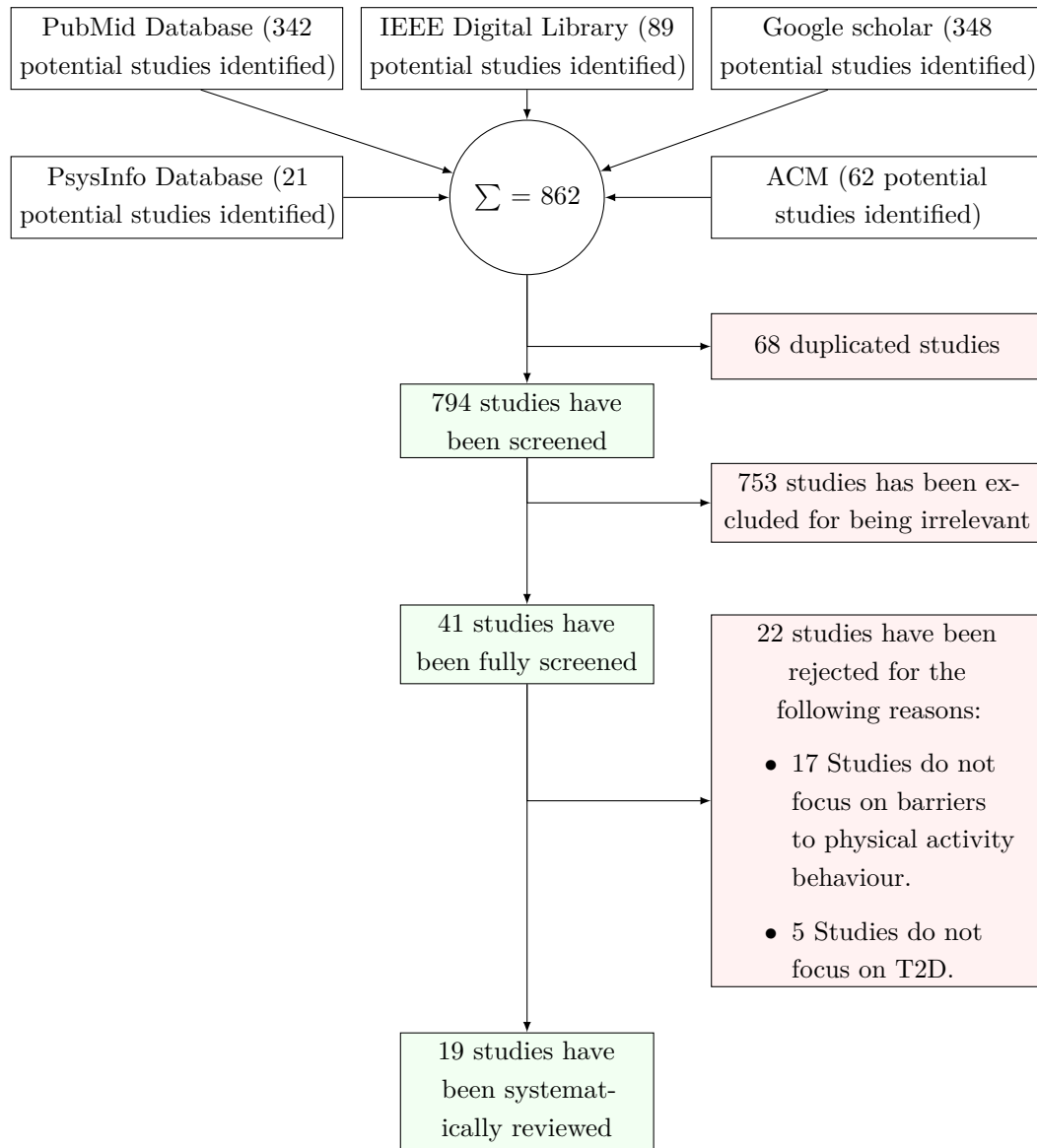


Figure 4.2: A flow-chart illustration of the Systematic Review.

Scoping Review. The systematic scoping review explores the identification of barriers to a specific behaviour (physical activity) for a specific demographic group (patients with T2D).

Results. Completion of the systematic review resulted into establishing a set of concepts and relations that were eligible for inclusion in our proposed ontology. Tables A.1, A.2 and A.3, in Appendix A, summarise the key results of this review. Table A.1, contains the discovered classifications that depict how the barriers would be hierarchically organised within the ontology. Table A.2 mainly reveals the barriers as discovered in the final reviewed studies, with instances in each row signifying a *prevents* relation between the identified barrier (Barrier column) and the corresponding physical activity (Physical Activity column). More specifically, barriers in the Barrier column were found to prevent activities in the “Physical Activity” column. Finally, Table A.3 shows instances as part of an *isSuggestedFor* relation (inverse relation of *prevents*). In other words, activities in Physical activity column were found to limit the effects of corresponding barriers in the barrier column.

Across all three tables is a “Source Quotation” column, in which we placed evidential quotes from a referenced study in support of the identified relations (Tables A.2 and A.3) and identified classifications (Table A.1) between the first and second column values. We give a deeper analysis of the tables in the following.

Table A.1 highlights and classifies the identified barriers. The identified barriers within the “Barrier” column are sub-classes of the classifications within the “Barrier Classification” column. To the right of the “Classification Barrier” column is the “Source Quotation” column, which contains a quote extracted from a referenced study as evidence to support the asserted classification.

The six main barrier classifications identified from the systematic review are: *health*, *physical*, *personal*, *environmental*, *social* and *psychological* barriers. In order to classify the identified barrier into one of the six barrier classifications, we use at least one reference. Examples to enhance readability of the table contents include, (1) “bad weather condition (e.g. cold)” which is classified as an environmental barrier [12, 22], (2) “feeling stressed or anxious” is classified as psychological barriers in [70, 240], (3) “lack of time” or “time restriction” is at different times classified as a personal barrier [12, 25] and an environmental barrier in [74]. We adopt the most commonly used classification for the “lackOfTime” barrier, which is personal barrier [12, 25].

For some of the barriers, for instance ‘obesity’ or ‘overweight’ [108] and ‘body image’ [203],

we could not find in the systematic review a reliable classification. For these, we relied on the evaluation study we detail in Chapter 5 (Section 5.3), by gathering domain expert opinions.

This classification process enabled us to build the hierarchical taxonomy of the Barrier Ontology, in Section 4.3.3.

Based on Table A.1, there are some individual barriers (“Barrier/ Classification” column of Table A.2) that can further be classified into sub barriers. For example, *weather condition* has been used as a barrier by [240] to embody different weather patterns, such as cold, hot [240] and rain [34]. In other scenarios, authors have explicitly mentioned factors indicate the presence of particular barriers, for example family obligations [172], children restrictions [75], and job commitments [108] were stated to indicate lack of time barrier.

Table A.2 contains identified barriers (barrier classifications and instances of these classifications), activities prevented by the barriers i.e. asserting a *prevents* relation, and proof of this asserted relation (quotes from sources annotated from the studies). An example to help in interpreting the table contents is row 3: Weather conditions (cold, hot, icy, raining) are barriers which prevent “Outdoor” activities (walking, cycling, jogging); (Misoon *et al.*, 2009). “The data were collected during the middle of winter when the weather is particularly cold in Korea, making it difficult to exercise outdoors” [240].

The most common barriers within the studies include bad weather [25, 108] or climatic condition [142], and time restriction [25, 143] or lack of time [75]. While barriers such as stress [25] and chest pain (e.g. breathing problems) [108, 159], are the least mentioned barriers.

Table A.3 captures activities suitable for managing or limiting barriers, specifically the inverse relation (*isSuggestionFor*) of the relation captured in Table A.2 (*prevents*). Activities under the “Alternative Suitable Physical Activities” column are discovered or suggested as a mechanism to limit the barriers in the activity column. An example to help interpret table contents is row 2: Indoor activities are suggested activities for Weather conditions barriers (hot and cold). (Misoon *et al.*, 2010) “Secondly, exercises that can be performed easily indoors should also be introduced as an alternative to outdoor exercise when it is too cold or hot” [240].

Conclusion. The systematic review, summarised in Table A.1, Table A.2 and Table A.3 not only contributes towards a fundamental part of this thesis, but also supports health workers to concurrently identify and classify barriers to behaviour. Although a few system-

atic reviews identify a barrier to physical activity for T2D, most reviews simply cite barrier identifications. Classification of barriers into similar groups supports a better understanding of the barrier behaviour domain; this essentially tries to address them collectively in groups instead of dealing with each single type of barrier independently. For example, a health barrier can be handled by health domain experts, while psychological barriers can be handled by health psychologists. Likewise, environmental barrier can be discussed with the responsible authorities (e.g. City Council).

4.3.3 Concepts of the Barrier Model

Section 4.2 defined the functional and non-functional requirements of the Barrier Ontology, and the systematic review (Section 4.3.2) allowed us to model the barrier concept by obtaining the barriers' terms and classifications; it allowed us to gain further understanding of the barrier domain by helping us to discover other related barrier concepts, and their relations, in order to limit these identified barriers. Therefore this model also includes the concepts that are needed not only to identify the barriers but also to suggest alternative activities to limit the barriers to physical activity behaviour for T2D. These concepts are: user (patient), physical activity, human disease (disease and current condition), BCT, stage of change, and belief. The person, physical activity, human disease (disease and current condition) and BCT concepts are all existing concepts that are imported into the model, which are clearly distinguished with a *green* background in Figure 4.1, whereas stage of change and belief concepts are built upon specific behaviour models. Each of these concepts is represented as a module that can be further specified individually. In other words, each concept classifies into self-contained modules that contain further details. This conceptual model is based on different research areas, such as e-health, psychological behaviour and physical activity. The concepts of the model are created and designed based on health-related behaviour theories and models. Each model concept and the relation among these concepts are discussed in detail below, starting with the related concepts. It is important to distinguish the notion of an individual concept from the conceptual barrier model; the concepts are each sub-elements of the Barrier Ontology this thesis proposes. We now describe how each concept has been created.

Barrier Concept

The barrier concept models the barrier to physical activity. As previously mentioned in Section 4.3.2, the six most common barrier categories, based on the systematic review are: *health, physical, environmental, psychological, and personal*. It's noteworthy to mention that, not all were simultaneously used in a single particular study, but that they were found across many studies. Health barriers pertain to the presence of a disease or condition, either temporary or permanent in nature; physical barriers pertain to the presence of a disability or temporary injury which affects physical capability; environmental barriers pertain to any natural environmental condition (e.g. weather) or any other factors affecting the environment (e.g. availability of local facilities or how safe an area is perceived); psychological Barriers pertain to a mental state or attitude of the person; and finally, personal barriers, related to all other aspects of the person's life, e.g. lack of free time, lack of support etc. Figure 4.3 presents the hierarchical classification or taxonomies of the six main barriers to physical activity based on the systematic review (Table A.1). Due to the difficulty of displaying the barrier's taxonomies (clearly) in one figure, each main barrier is independently (Figures 4.4, 4.5, 4.6, 4.7, 4.8 and 4.9) illustrated to clearly display its classification decomposition or hierarchy. This barrier's taxonomies are transformed into an ontology in Section 4.4, in order to create the barrier concept within our conceptual barrier model.

Physical Activity Concept

In order to model or present the physical activity concept in the conceptual barrier model, the existing physical activity ontology³ (Figure 4.10) is imported. The physical activity ontology aims not only to decide the type of activities that are prevented by the identified barriers, but also to suggest the suitable alternative types of physical activity to limit these identified barriers.

The physical activity ontology is imported from the National Centre For Biomedical Ontology (NCBO) BioPortal Ontologies⁴, one of the most widely used ontology repositories for health applications. This ontology is open-source and free to use. It includes 29 classes, 37 object properties and 243 axiom (e.g. data properties to define classes). This ontology classifies the type of physical activities as a hierarchy. For example, exercise

³<http://bioportal.bioontology.org/ontologies/SMASH?p=classes&conceptid=root>

⁴<https://www.bioontology.org/>

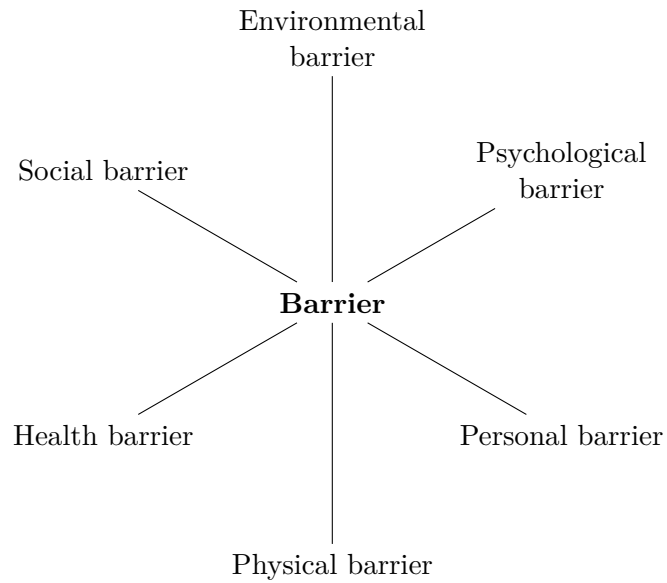


Figure 4.3: Main barriers' classifications.

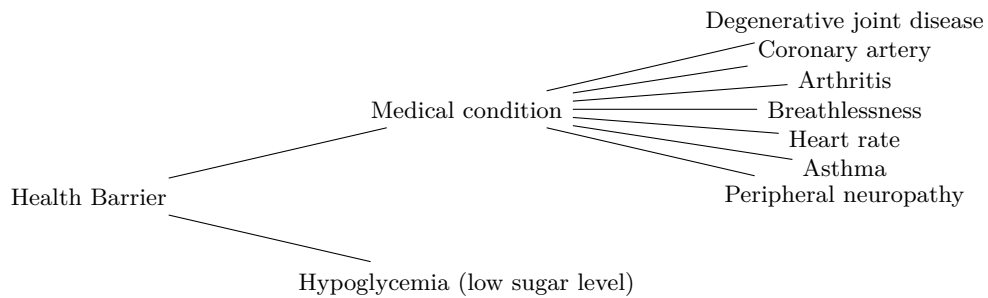


Figure 4.4: Hierarchical Health barrier classification (a continuation from Figure 4.3).

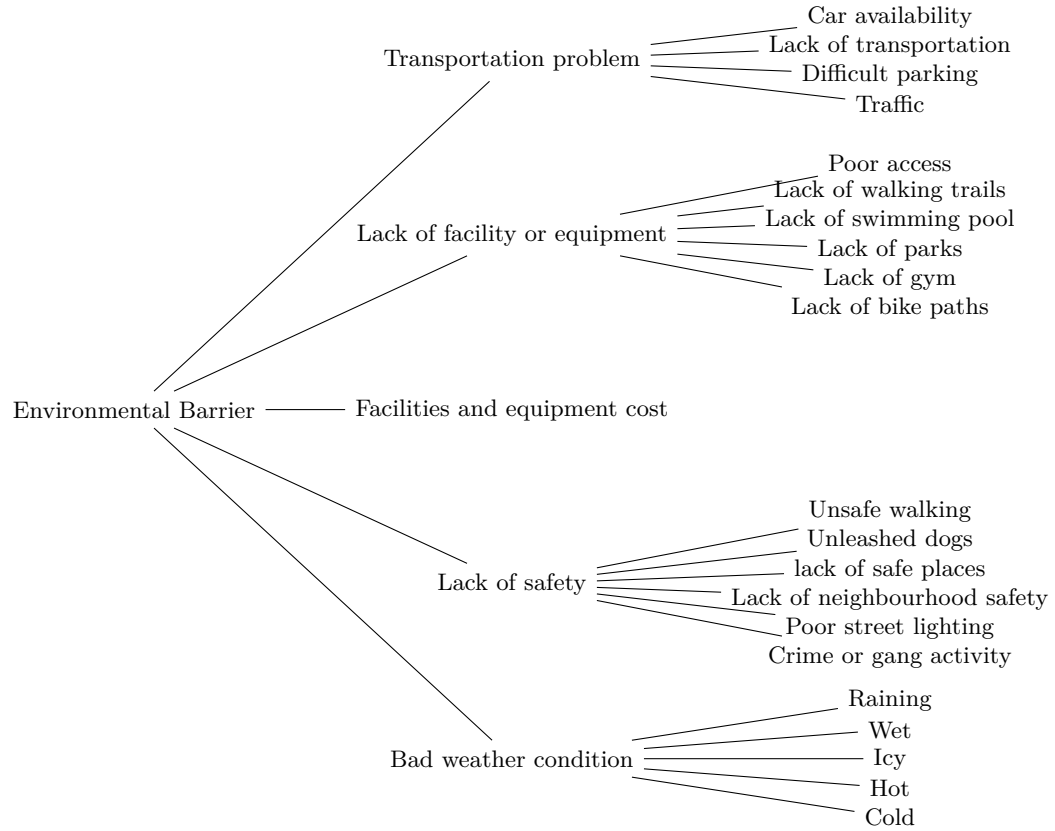


Figure 4.5: Hierarchical Environmental barrier classification (a continuation from Fig. 4.3).

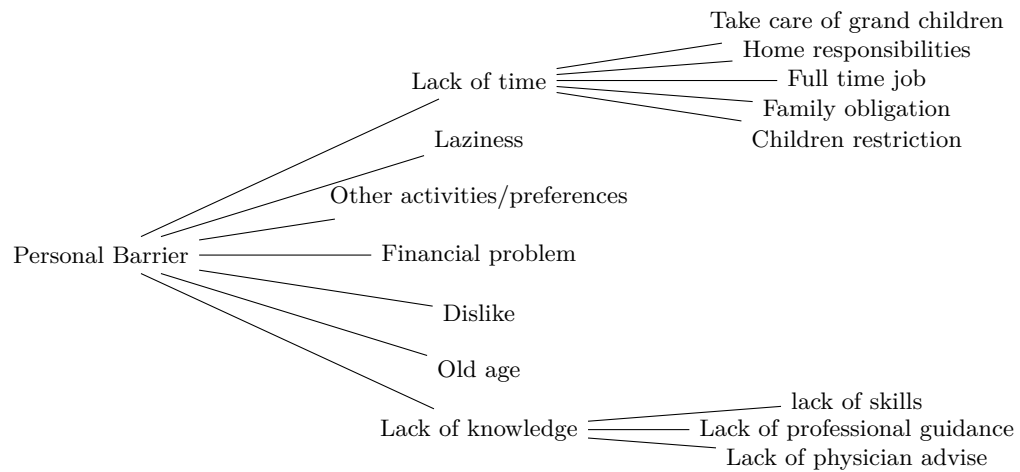


Figure 4.6: Hierarchical Personal barrier classification (a continuation from Figure 4.3).

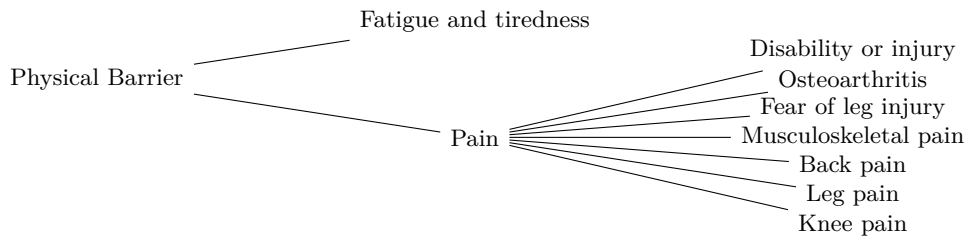


Figure 4.7: Hierarchical Physical barrier classification (a continuation from Figure 4.3).

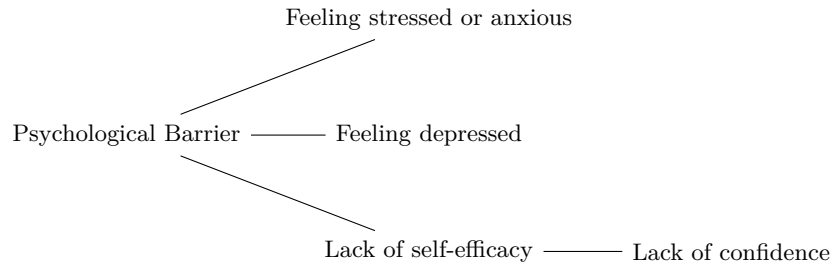


Figure 4.8: Hierarchical Psychological barrier classification (a continuation from Figure 4.3).

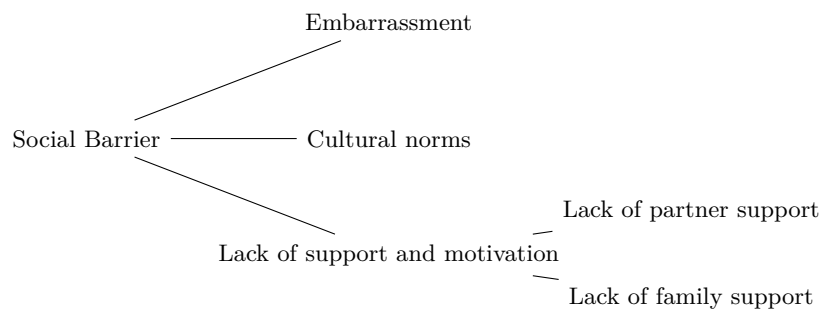


Figure 4.9: Hierarchical Social barrier classification (a continuation from Figure 4.3).

classifies into three main categories: aerobic exercise, anaerobic exercise, and flexibility exercise. Furthermore, aerobic exercise is then divided into running and walking.

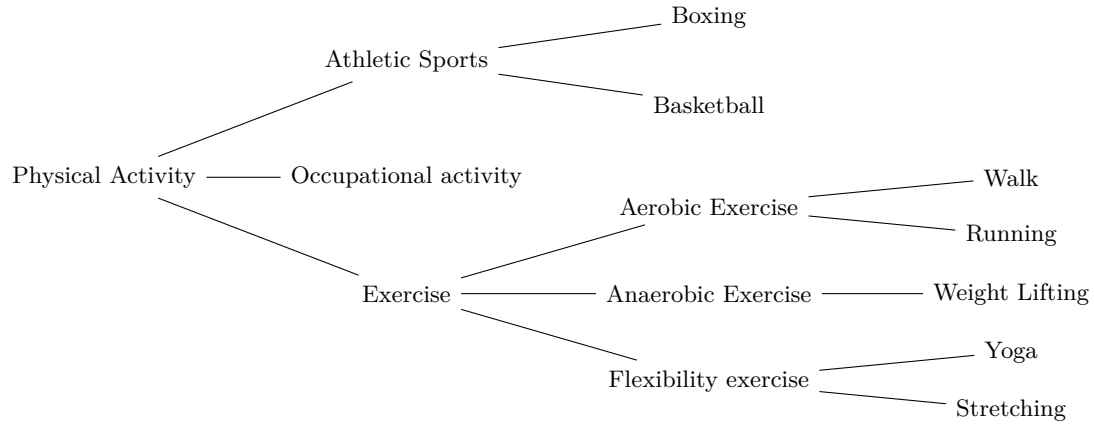


Figure 4.10: Hierarchical classification of physical activity ontology (partial view).

The physical activity ontology is an expandable ontology, so the physical activity ontology is extended to include more types of physical activity, based on specialized studies such as the Compendium of Physical Activities⁵ [2], to better meet the conceptual model's objective of finding an appropriate physical activity based on identified barriers. For example, home activities such as cleaning, household tasks and carrying things upstairs are added to the physical activity ontology, in order to limit environmental barriers such as bad weather condition (e.g. cold or hot) [240]. Additionally, by referring to the Compendium of Physical Activities, the Anaerobic Exercise (Weight Lifting) and Flexibility Exercise (Stretching and Yoga) instances from the physical activity ontology are merged into one, known as Conditioning Exercise. Figure 4.10 shows classification of some physical activities, based on extension of the physical activity ontology and study of the Compendium of Physical activities.

Other ontologies, related to exercises (not daily activities) are exercise ontology and the ontology of physical exercises which are discussed in Section 3.8.7. Based upon our review of ontologies from common ontology repositories, we can conclude that the physical activity ontology from BioPortal is the only existing ontology that met the conceptual model's needs to model the physical activity concept.

⁵<https://sites.google.com/site/compendiumofphysicalactivities/Activity-Categories>

Human Disease Concept

The human disease concept is aimed at not only identifying more health (diseases) barriers, but also for identifying more psychological barriers such as ‘anxiety’. So, the existing disease ontology⁶ is imported into the barrier model to achieve the human disease concept’s goals, rather than creating one from the scratch. The disease ontology looks to present each type of health disease (barrier) based on etiological classification and to annotate biomedical data across different biomedical resources, such as NCI, the NCI thesaurus (NCIt), MeSH (which is extracted from the Unified Medical Language System (UMLS) with mapping to SNOMED CT), and the International Classification of Diseases (ICD) of the WHO [205]. Consequently, the number of diseases identified has now risen to 46,000, along with increased amounts of medical terms and the capability to share knowledge with other resources [131, 230].

The disease ontology presents both physical diseases, and mental (psychological) health conditions [189], which classify as a type of disease. The domain of this ontology is disease, and the ontology categorises into eight main groups: disease by infectious agent, disease of anatomical entity, disease of cellular proliferation, disease of mental health, disease of metabolism, genetic disease, physical disorder, and syndrome (Figure 4.11). Each group includes further classification to provide a clear definition for each of their diseases. For example, the ‘disease of metabolism’ group classifies into acquired metabolic disease and inherited metabolic disease; the acquired metabolic disease classifies into amyloidosis, carbohydrate metabolism disease, and so on, until the condition of T2D is reached, as demonstrated in Figure 4.11. The disease ontology is open-source and free to download from two different ontology repositories: the BioPortal⁷, and the Open Biomedical Ontology (OBO) Foundry⁸. The disease ontology is designed using OWL, and is considered a well evaluated and reliable ontology [230].

⁶<http://disease-ontology.org/>

⁷<https://biportal.bioontology.org/ontologies/DOID>

⁸<http://www.obofoundry.org/ontology/doid.html>

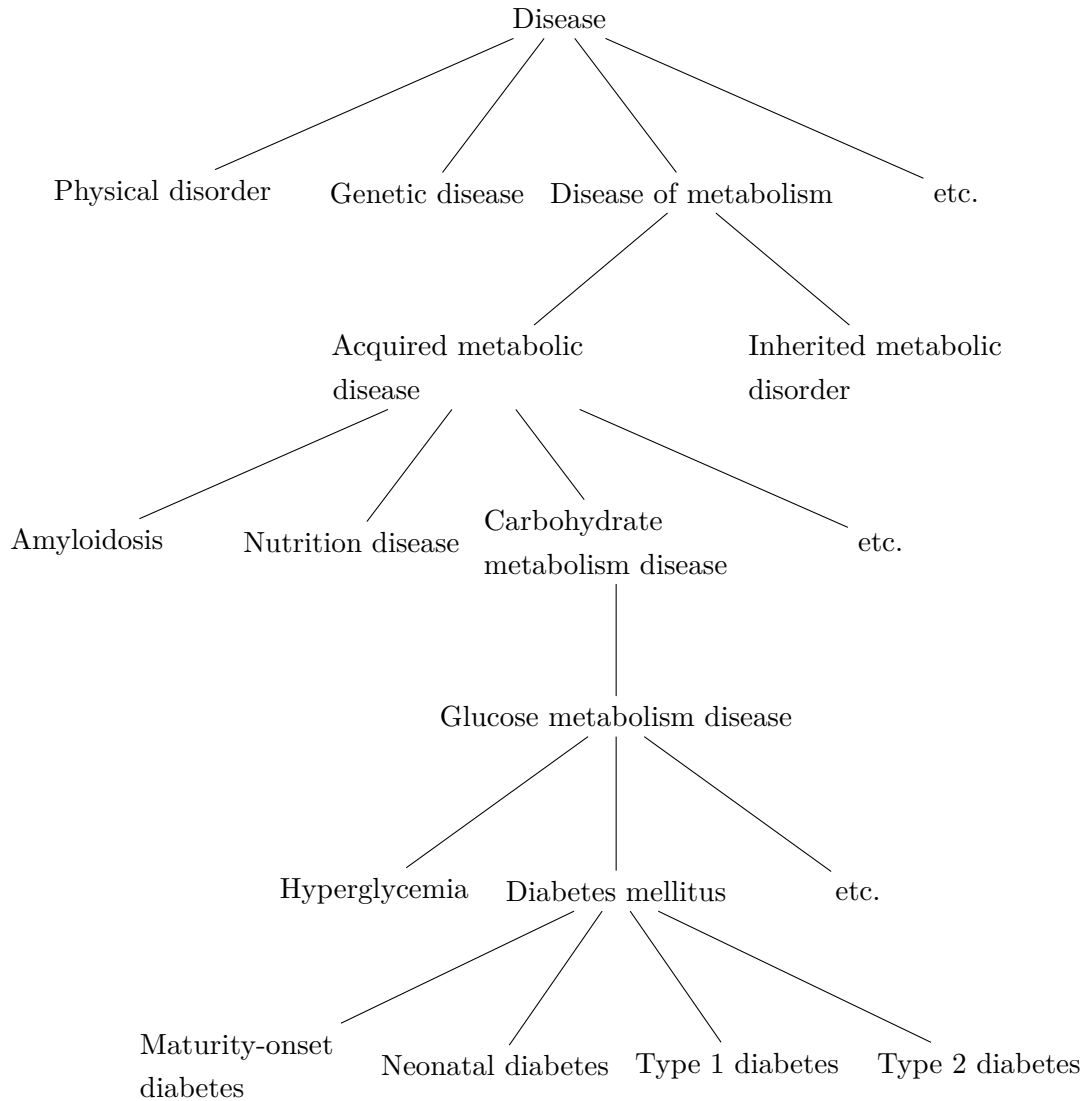


Figure 4.11: Partial view from the disease ontology (identifying T2D as a disease)⁹.

The human disease ontology updated version (1.2) that is imported into the Barrier Ontology was released on 18/04/2019 (Section 4.4 (step of import existing ontology)). It includes a 13285 Class count, 13 object property count, and 138704 axioms. So, the fact that we can summarize the updated version by identifying both health and mental disease and as well integrate disease ontology with a different source, such as NCI, essentially backs

⁹<http://disease-ontology.org/>

up or rather supports our choice to import the disease ontology into the Barrier Ontology. Other examples of disease ontology are the ICD of the WHO [205] and human phenotype ontology [135].

Figure 4.13 demonstrates part of modelling activities in the physical activity ontology, imported into to the conceptual model using the Protégé 5.0.0 ontology editor (Section 4.4).

General User Model Ontology

The General User Model Ontology (GUMO) aims to model ideas, entities, and relations that are primarily related to the system user, or the patient. In order to present the user concept, the existing ontology of General User Model Ontology (GUMO) [105] is imported into the conceptual barrier model. The GUMO is based on User Models (UMs), which aim to facilitate the reuse of user modelling data, requiring semantics to enrich the ontology representation. User Models were first discussed, in 1990 [125] and were later developed in 2003 [27, 219]. In the conceptual barrier model, the GUMO is utilised to model the patient concept, which includes characteristics modelled as sub-classes. These include: personality, patient profile such as name, contact information, patient record number or ID number, job/employment status, address, and date of birth. Some of the patient's terms are further classified into sub-concepts (e.g. patient address).

Conceptually, the GUMO is designed based on the USERML approach [104, 105], which divides user model dimensions into three main parts of situational statements which are auxiliary, predicate and range or probability. For example if the patient is interested in a particular type of psychical activity (e.g. running), the auxiliary is the relation (*hasInterest*), the predicate is the activity type (*running*), and the range is the activity intensity (huge, moderate or low). As a result of a massive group of auxiliaries, predicates and ranges, the authors realised that everything would potentially be modelled as a predicate for auxiliary names; *hasInterest* and *hasKnowledge*. They therefore suggested a solution to only identify basic user model dimensions.

In our conceptual barrier model, we adopt GUMO to meets our requirements. Utilizing published work from studies [105–107], we study the representation of the various user dimension models and build an ontology that largely imitates GUMO. Despite its online unavailability (with several failed/corrupted domains and links to the complete ontology

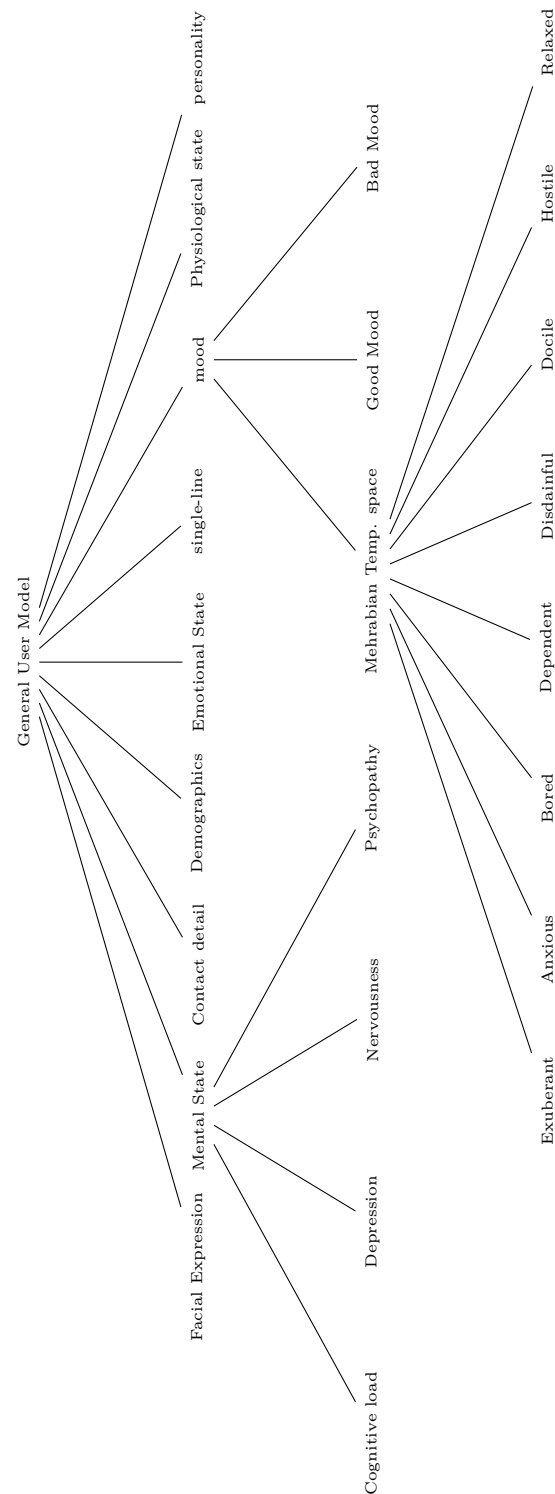


Figure 4.12: Partial view from the General User Model Ontology (GUMO) [105].

such as¹⁰ and¹¹ [105]), we compile information from the above mentioned sources under each of the listed modelled basic user dimensions. For example Section 5.2.2.3 in [107] indicates that Mood (a basic user dimension) can be good or bad or Mehrabian Temperament Space (Exuberant, Dependant, Relaxed, Docile, Bored, Disdainful, Anxious and Hostile). The GUMO, which is based on the Barrier Ontology's requirements, includes 211 classes and 646 axioms. Figure 4.12 also illustrates a basic GUMO concept. Figure 4.1 demonstrates how the GUMO links to stage of change (Section 4.3.3) and belief (Section 4.3.3) concepts through *hasStageOfChange* and *hasBelief* relations, respectively.

There exists several other ontologies that encapsulate wider aspects of User (human) activities i.e. better referred to as User Modelling Ontologies. Nevertheless, GUMO stands out mainly because of its fluidity when exchanging user model data across different user-adaptive systems [106], because it adopts a global semantic web language OWL [63]. Additionally, other GUMO extensions such as User Navigation Ontology (UNO) [132] constrain user-modelling to just a single activity which is Navigation and wayfinding in this case. On the other hand, user profile ontologies such as OntoPIM (Ontology Personal Information Management) describe various users' dimensions and share a lot of concepts with GUMO. OntoPIM describes the user's domain of interest covering user characteristics such as personal information, general user characteristics, user abilities, preferences (i.e. abstract likes/dislikes), interests, activities and profession, which are rather static information [153]. Nonetheless, the latter characteristics describe only the user profile and not the user context, such as a user current position which GUMO does in its Role dimension.

Behaviour Change Technique Taxonomy

The Behaviour Change Technique taxonomy (BCT) is not a formal, machine readable ontology, as there is currently no ontology explicitly for behaviour change, however BCT is an existing effort aiming to create a taxonomy of barriers and behaviour change techniques. We look forward to seeing the BCT as an ontology in the future, to add it to this research. The BCT is discussed in further detail in Section 3.8.1. It is important to note that the lack of a machine-readable ontology that captures BCT does not negate the importance of the conceptual barrier model presented in this thesis, either as barrier concept by itself (requirements 4 and 5 in Section 4.2), or when integrated with other existing ontologies to

¹⁰<http://www.gumo.org/>

¹¹<http://www.u2m.org/>

form the Barrier Ontology (to achieve requirements 1, 3 and 4 from Section 4.2).

Stage of Change

The stage of change concept is based on the TTM, and is sometimes known as the “stage of change theory” (Section 2.2.2). This concept include the following terms/vocabularies: pre-contemplation, preparation, and action or maintenance. All possible current patient behaviour is represented or identified by one of these terms. This makes it possible to determine the appropriate type of interventions (e.g. suggestion) to change and maintain the patient’s behaviour.

Belief

The belief concept includes the terms ‘work’, ‘health’ and ‘society’s belief’, which all have a high priority on the life of the patient. These terms play an essential role in influencing the patient’s current behaviour towards change or behaviour maintenance. For example, a patient who does not engage in physical activity (pre-contemplation) but has a high priority belief in the health, would benefit from advice such as regular physical activity to prevent disease and associated complications (e.g. T2D), which could influence change in the patient’s currently inactive behaviour.

4.3.4 Relations of the Barrier Model

This section discusses the relations among the concepts of the conceptual barrier model (Figure 4.1). We focus on *prevents*, *isSuggestedFor*, *subClassOf*, and *BarrierSign* relations - all of which either are direct relations of the barrier concept, or directly reflect the ontology’s requirements (Section 4.2).

As discussed in Chapter 3 (Section 3.6), the OWL language chosen to model the Barrier Ontology distinguishes between two different types of properties in modelling ontologies i.e. Object property and Datatype property relations [48, 50, 192].

The former serves to link instances of different classes, whereas the latter is used to assign class instances values or literals. Besides these, we also present hierarchica; relations in the rdf/owl schema notation that are responsible for properties such as *subClassOf*.

Object Property Relations

***prevents* relation.** The *prevents* relation is one of the main relations in the barrier domain, linking the barrier concept with physical activity concept. This relation aims to identify the barrier that prevents patients from engaging or participating in specific types of physical activities. Therefore, the barrier concept is the domain of this relation, and the physical activity concept is the range of this relation. This relation serves to answer queries such as ‘what are the types of barrier that prevent a patient from performing outdoor walking?’, hence fulfilling requirement (2) in the requirement specification (Section 4.2).

The *prevents* relation is clarified in Table A.2, which is based on the systematic review (Section 4.3.2). Table A.2 shows that some barriers (listed in the Barrier/Classification column) are linked to specific types of physical activity (listed in the Prevented Physical Activities column), such as outdoor activities that link with walking activity, whereas other barriers are not linked to particular activities, being considered as barriers to physical activity in general (i.e. no specific type of activities are included). This can be interpreted as the barrier negatively affecting any physical activity. For example, the ‘lack of time’ barrier can exert negative influence on many different types of activities, such as e.g. walking and swimming [142]. The question of whether or not barriers prevent activities (either specific activities, or activities in general) is later evaluated by domain experts using the CQs technique in Chapter 5 (Section 5.3).

***isSuggestedFor* relation.** The *isSuggestedFor* relation is the inverse of the *prevents* relation, meaning the *isSuggestedFor* relation is restricted with the *physical activity* concept as domain, and with the *barrier* concept as range. This relation aims to link barriers to suitable alternative activities, in order to limit the identified barriers. Table A.3, which forms part of the systematic review, presents barriers and suggested alternative activities. The *isSuggestedFor* relation aids answering research sub-question (c); ‘How can we model the Barrier Ontology to suitably recommend or suggest solutions to barriers to physical activity for T2D patients?’, by using a CQ such as ‘What physical activity is suggested for the bad weather condition (‘*raining*’) barrier?’ (Section 5.3.2). This in turn satisfies requirement (3) of the requirement specification (Section 4.2).

N-ary Relations

diseaseConditionRelation . The relation of *diseaseConditionRelation* requires further explanation; this relation is an n-ary relation¹² [23, 194], considered as a container to hold additional or temporal descriptions about a specific instance of the patient concepts. The main aim in creating this type of relation is to allow representation of a specific disease and current condition for a specific patient. This means it is possible to present a case of the form: “patient ‘X’ has disease ‘Y’ and currently has condition ‘Z’” (e.g. Patient Smith has asthma and currently has fever). The three arguments of the original n-ary relation *patient*, *disease*, and *CurrentCondition* leads to three true binary relationships: *has_disease*, *disease_value* and *currentCondition_value* [262, 267]. Figure 4.1 demonstrates how the *diseaseConditionRelation* is related to barrier concept (health barrier) through the *isTriggeredBy* relation.

rdfs:subClassOf Relations

The *subClassOf* relation (shown in Tables A.1 and A.2) implies that instances belonging to one class, inherently also belong to the parent class. It often depicts a parent-child relationship between a class and its sub-classes. For example, *LackOfTime* holds a *subClassOf* relationship with the *Personal* barrier; similarly, *Pain* holds a *subClassOf* relationship with *Physical* barrier; this implies that the ‘*LackOfTime*’ and ‘*Pain*’ are subclasses of *Personal* and *Physical* barriers respectively. By default, instances belonging to the subclasses also belong to the parent classes, yet the inverse doesn’t hold true; if ‘*Pain*’ is a *subClassOf* ‘*Physical Barrier*’, then ‘*Physical Barrier*’ cannot be a *subClassOf* ‘*Pain*’. The *subClassOf* relation is essential in answering CQs such as ‘what is the barrier type of the ‘lack of time’ barrier?’ and ‘which barriers are classified as a ‘personal’ barrier?’, and in satisfying requirements (4) and (5) of the requirements specification in Section 4.2.

Datatype Property Relations

***barrierSign* relations.** The systematic review (Section 4.3.2) not only aims to obtain the necessary knowledge (e.g. related barrier’s terms and relations), but also provides an understating of the barrier’s domain. Some barriers in the barrier class can be logically inferred or easily predicted based on the user’s profile information. For example, barriers

¹²<https://www.w3.org/TR/swbp-n-aryRelations/>

such as ‘old age’ and ‘lack of time’ are predictable, given the user characteristics of age and job/employment status, respectively. Other barriers, such as medical condition (e.g. heart attack), fatigue, and anxiety are easily deduced from the user’s health profile.

The *barrierSigns* are DataType Properties (DTP), each of which can be used to assign a value (such as a string or integer) to an individual of the type or domain class Patient. Each barrierSign can independently link the user to barrier concepts. For example, the barrierSign DTP employmentStatus takes on a value such as "hasFullTimeJob", and can be annotated with a utility property “seeAlso” [39, 58] which links it to the barrier class “LackOfTime”.

These sub DTPs under the *barrierSign* relation are vital in enabling prediction of a specific barrier on the basis of information captured in the user classes. The *barrierSign* object matches sub-classes of the Barrier class to sub-classes of the Patient class, hence is restricted to the domain of Patient. The *barrierSign* relation contributes to satisfying requirement (1) of the requirement specification (Section 4.2), and therefore answering CQs such as ‘what is the expected barrier of a user whose employment status is “Full Time Job” and is suffering from “Asthma” disease?’.

4.4 The Ontology Development Process

We now discuss how the conceptual model of barriers to physical activity behaviour is developed using the “Ontology Development 101” and using the Protégé editor (Section 3.6). Justification for the “Ontology Development 101” [192] methodology is presented in Chapter 3 (Section 3.6). As mentioned in (Section 3.6), it is not necessary to always follow a design methodology, such as the 101 methodology or those presented in [77, 242, 253]. However, it is often advisable to do so in order to ensure that the steps of ontology development are complete. Another benefit of using a prescribed development process is that a well narrated development process enhances review and re-use of the ontology, aiding researchers keen on re-producing similar or related work.

In order to develop the Barrier Ontology, the ontology editor tool Protégé (Section 3.6) is used. It is important to note that the term ‘*concept*’ of the conceptual model is replaced with the term ‘*class*’, which is more appropriate during the development process of Barrier Ontology (Section 4.3). We now detail below the development steps that are used to develop the Barrier Ontology.

1. Decide the domain and scope of the Barrier Ontology:

Physical activity for T2D is chosen as the domain and scope of the Barrier Ontology. The reasons for this are to define the domain and scope of ontology (i.e. - the ability of physical activity to manage T2D - one of the most common diseases worldwide), and to put the problem of behaviour change (barrier) into context, as discussed in Section 4.3.1. Different types of CQs (Section 3.6) are defined to further limit the ontology's domain and scope (barrier to physical activity for T2D), and to ensure the ontology's requirements (detailed in Section 4.2) are entirely covered. These CQs are divided into four templates to meet the Barrier Ontology's requirements (Section 4.2). Those presented here primarily focus on the barrier concept, however, they are transferable across other concepts such as physical activity and GUMO. These CQs are also relevant in evaluating the Barrier Ontology in Chapter 5. Below is an example of each of these four templates. More examples and details are presented during the evaluation of the Barrier Ontology (Section 5.3).

- What physical activity is suggested for the bad weather condition (*'raining'*) barrier?
 - What barriers prevent or limit people from performing 'Football', despite being interested in this activity?
 - What is the expected barrier for a male patient who has a *full-time job*, lives with *his wife* and *two children*, and complains about *not having friends* to play sports with, and *no personal support* for performing physical activity?
 - What is the barrier type of *'transportation problem'* barrier?
 - What barriers are classified as *'environmental'* barrier?
2. Consider reusing existing ontologies:

Reusing existing ontologies, instead of creating one from scratch, is a good practice and a powerful process in ontology development [113, 192]. The step of enumerating related terms comes after this step, in order to ensure there is no existing ontology or hierarchical taxonomy to reuse or to extend into the given domain.

Four existing ontologies are imported into the Barrier Ontology to model some concepts. These four ontologies are the human disease ontology¹³, physical activity ontology¹⁴, GUMO [105] and BCT [168]. As mentioned in Section 4.3.3, there is currently no BCT ontology or taxonomy that can be imported into the Barrier Ontology, so BCT was expressed as an empty class. Figure 4.13 presents physical

¹³<http://disease-ontology.org/>

¹⁴<http://bioportal.bioontology.org/ontologies/SMASH?p=classes&conceptid=root>

activity concept in the conceptual barrier model.

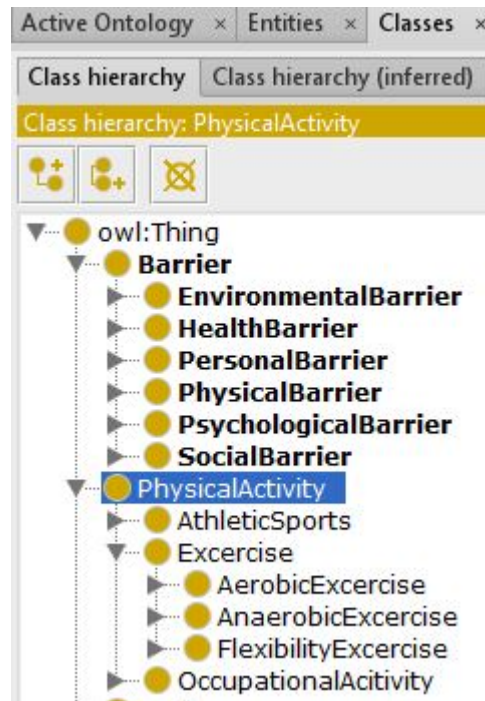


Figure 4.13: Physical activity ontology imported into the conceptual barrier model.

3. Enumerate important terms in the barriers ontology:

The step of enumerating related terms comes after ensuring there is no existing useful ontology or hierarchical taxonomy that can be reused or to extended into the Barrier Ontology, instead of creating a new ontology. So, this step includes classes that are created based on specific models (belief and stage of change) and class that are created from scratch (barrier class). Regarding the barrier class, a systematic review (Section 4.3.2) is achieved to enumerate the important terms relevant to our main concepts i.e. barrier and physical activities.

4. Define the classes and the class hierarchy:

This step aims to refine the list of terms from the previous step (enumerate important terms), that can be classified as classes and subclasses. The Top-down method (Section 3.6), which starts with the most general class, progressing to the specific class in the domain, is used to create class hierarchies, thus instantiating a taxonomic structure. The hierarchical taxonomies among the barriers terms help us achieve

requirements (4) and (5) in the requirements specification (Section 4.2), as well as answer the CQs such as ‘What is the barrier type of ‘*transportation problem*’ barrier?’ and ‘What barriers are classified as ‘*environmental*’ barrier?’. In this step, we also define the extended classes and the class hierarchy of the extension to the physical activity ontology. It is important to note that the classes and the class hierarchy of the existing ontologies, such as disease ontology and GUMO, have already been processed automatically during the ‘importation of existing ontologies’ step (detailed in Step 2). Additionally, this step includes modelling the stages of change (Section 2.2) and belief (Section 4.3.3) concepts, hierarchically.

5. Define the properties of classes:

In order to infer or predict a barrier, with respect to a given user characteristic, it’s essential to model a relationship between the user concept and the barrier concept. For purposes of modelling this relationship, GUMO is extended to include Data Type Properties (DTPs) that link user classes to the barrier classes. We use the DTP relations (specifically we use sub DTPs of the top-level data property relation *barrierSign*, as introduced in Section 4.3.4). Each data property in GUMO is created as an object belonging to the Patient domain class and is annotated with a comment (literal/string value) and a utility property instance “*rdfs:seeAlso*” [39]. The “*rdfs:seeAlso*” instance links the DTP to a barrier class. For example, the DTP entry “*hasFullTimeJob*” (a sub-property of *barrierSign*) can be annotated with a literal value “yes” and as a barrier class “*LackOfTime*” using the utility property *rdfs:seeAlso* (see example in excerpt below).

rdfs:seeAlso is an instance of rdf:Property that is used to indicate a resource that might provide additional information about the subject resource [39].
The rdfs:domain of rdfs:seeAlso is rdfs:Resource (“hasFullTimeJob” data property).
The rdfs:range of rdfs:seeAlso is rdfs:Resource (“LackOfTime” barrier class).

This initialized relationship carries both a conceptual meaning and logical meaning. A conceptual meaning implies that there exists a relationship between the employment status DTP “*hasFullTimeJob*” and the “*LackOfTime*” barrier class; a logical meaning implies that “*hasFullTimeJob*” is a sign which indicates the presence of the “*LackOfTime*” barrier.

Another example of this linking of a DTP and barrier class can be found with the

DTP “isSocial”, which can be annotated with a literal value “yes” and with the barrier “LackOfSupportAndMotivation”. Further examples include the DTPs “age”, “hasIllness”, “hasT2D”, “hasOtherDiseases”, as well as others induced from the domain knowledge gathered. These are further detailed in Table A.2.

6. Define the restriction of the classes:

This step includes setting and applying the restriction on object properties relations (Section 4.3.4) among the classes of the conceptual model. These restrictions describe the range and domain for each of the domain classes. For example, the *prevents* relation is restricted to the domain *barrier*, and range *physical activity*.

7. Create instances:

The method used to determine the instances, described in step 4 of the development proves (define the classes and the class hierarchy), is a ‘top-down’ method which starts from the general classes and progresses to the precise classes, with the finest granularity of classes presented as instances. Instances are therefore defined as “the most specific concepts represented in a knowledge base” [192]. So, the final description of the barrier terms in Figures 4.4, 4.5, 4.6, 4.7, 4.8 and 4.9 are created as instances in the barrier class.

In order to complete this step, the following stages should be followed: (1) choose a particular class (Step 4); (2) create the required instances of the selected class; (3) instantiate the necessary properties to capture the relations for the instance. For example: (1) the ‘HomeActivity’ subclass is chosen; (2) ‘Cleaning’, ‘Vacuuming’, ‘RopeSkipping’, etc. are created and specified as instances belonging to the ‘HomeActivity’ subclass; (3) the relation (*isSuggestedFor*) is instantiated to link these three instances to the weather condition instance “cold”, which literally means cleaning, vacuuming and rope skipping are ideal alternative activities for the barrier of cold weather.

This implies that sub-question (c) of the thesis “Can we use the ontology to identify the best suggestion to give to a subject with T2D in terms of types of physical activities to limit some perceived or objective barrier?” is answered.

A further example is if the inverse relation *prevents* is used: (1) ‘LackOfTime’ is chosen as the subclass; (2) ‘FamilyObligations’ and ‘HomeResponsibilities’ are instances created under the ‘LackOfTime’ subclass; (3) the *prevents* property is used to link these instances to outdoor activities instances such as ‘Walking’ and ‘Running’.

4.5 Summary

This chapter narrates the various steps undertaken to design and develop the Barrier Ontology. It establishes an understanding of the various decisions and steps taken in developing an ontology. Sufficient explanation of the domain terminologies assists the reader to form an intuitive appreciation of the domain, particularly its relevance to the research. At its core, it makes use of the “Ontology Development 101” process to guide the development of the Barrier Ontology. It started with selection of a domain and scope, followed by defining a set of CQs to extract terms and concepts necessary in building the hierarchies of the ontology. Completion of the ontology construction answered sub-question (a) *“How can we build an ontology of “barriers to physical activity for T2D patients”?”*.

After a comprehensive literature review, three different ontologies are deemed fit for integration into the Barrier Ontology: disease ontology, physical activity ontology, and GUMO. Successful re-use of existing ontologies directly answered sub-question (b) *“How can we demonstrate the use of a formal methodology, including the notion of ontology reuse, to objectively support the Barrier Ontology?”*. Some relations included in the ontology, particularly, the *isSuggestedFor* relation, indirectly implies an activity is suitable or recommended for limiting a certain barrier; this answers sub-question (c): *“How can we use the Barrier Ontology to produce suitable recommendations of physical activities which take into account barriers to such physical activity from T2D patients?”*. This subquestion is further addressed in the evaluation study, as detailed later.

The Barrier Ontology developed in this chapter is empirically evaluated in the next chapter (Chapter 5), making use of an analytical data-based technique and a competency question-based approach.

Chapter 5

Evaluation

5.1 Introduction

Chapter 4 presented the process and methods used in construction of the Barrier Ontology, uncovering how the barrier conceptual model was developed, and how the ‘Ontology Development 101’ development process was used in building the Barrier Ontology. This chapter evaluates the ontology deliverable from Chapter 4.

There is currently no consensus on a global approach or methodology for evaluating ontologies [84], however there is research which focuses on structural aspects, functional aspects, and usability aspects, with regards to ontology validation and verification [67, 196]. This chapter details two methodologies that were systematically adopted to evaluate the Barrier Ontology, specifically data-driven evaluation [38] and competency questions (CQs) evaluation [99, 193, 261]. Data-driven evaluation (Section 5.2) involves comparing an ontology with a corpus of texts extracted from sources relevant to the ontology knowledge domain [38], for our case, these sources may include articles in open access medical repositories such as PubMed¹, and are described in further detail in Section 5.2. The objective of the comparison is to identify co-occurrences of the terms and relations between the ontology and the corpus. Data-driven evaluation employs metrics such as *precision* and *recall* to measure the lexical key word coverage of the ontology, and hence its overall fit to the domain [38]. In Information retrieval (IR), *precision* is the proportion of all items retrieved that are relevant or correct, and *recall* is the proportion of all relevant or correct items that are retrieved [42].

¹<https://pubmed.ncbi.nlm.nih.gov/>

Brewster *et al.* [38] discuss the challenges in identifying relations in the data-driven approach, indicating that despite relations such as *isA* (hyponymy) and Part-of (meronymy) being common, it is not clear that they are of the right granularity to represent knowledge in the ontology. We leverage on the competency questions (CQs) approach (Section 5.3) and knowledge gathered in a survey from domain experts (such as health practitioners and others with T2D knowledge) to evaluate the ontology’s ability to correctly address or answer the CQs, as well as to evaluate the relations in the ontology as briefly discussed here and expanded on in Section 5.3.2. CQs are designed with pre-defined answers that an ontology has to correctly identify [99], by means of returned answers to queries. We use the survey gathered answers to the CQs as the pre-defined answers the ontology has to correctly identify. Therefore, the CQ based evaluation objectives are achievable mainly by comparing the ontology retrievable answers to survey gathered answers for a set of CQs. Positive correlation or high similarity between the answers to the CQs from the two answer-bases (ontology and survey) imply that the ontology correctly addresses CQs and therefore provides formal justification of the ontology’s competence and ability to satisfy its intended use [30, 99, 193]. Constructing queries that contained the correct relations (which is vital in IR [47]) allowed us retrieve answers (i.e. domain or range of the respective relations) to the CQs. Subsequently, this enabled us to examine the precision of the domain or range of the candidate relations [91] (entailed in the CQs) as well as to determine how comparable the answers were to those feedback gathered from domain experts through a survey. Section 5.3 describes in detail the CQs as well as survey gathered knowledge used for evaluation.

5.2 Data-driven Approach

Brewster *et al.* [38] suggested using a data-driven approach to evaluate the degree of structural fit between an ontology and a corpus of documents. Unlike the golden standard approach, which requires an existing ontology [151], the data-driven approach broadens the scope of the knowledge base. It allows one to utilise multiple data sources in deciding if an ontology meets all its requirements and is a good representation of a particular domain of knowledge. We use a data-driven approach to evaluate the terms and concepts in the Barrier Ontology against a corpus of text extracted from articles, clinical trial systematic reviews and medical reports in medical journal sites and open access biomedical databases

such as PubMed², British Journal of General Practice (BJGP)³ and ClinicalTrials.gov⁴. The sources (included in Appendix B.2) were obtained by retaining a subset of numerous google search retrieved results that were relevant to the domain. The steps covered in this evaluation are illustrated in Figure 5.1.

5.2.1 Comparison of the Barrier Ontology with Documents from the Domain

In using the data-driven approach, the objective is to compare content at the various levels in the ontology with the content of the corpus, thereby identifying similarities and differences (in terms of existing concepts) between the two. This process is vital in measuring the lexical key coverage that the ontology attains on the corpus, i.e. how many key words (concepts) co-occur in the ontology and the corpus. This quantifies the extent to which the ontology reflects the domain from which the corpus was extracted. Strictly speaking, the data-driven approach requires counting the number of overlapping terms or concepts between the ontology and the corpus. A 4-phased approach is employed during the comparison process. These phases include text summarisation techniques, such as Topic modelling (phase 3), in order to enable efficient processing of the corpus. Ideally, these techniques extract a small group of entities (terms) that can be individually compared with the terms and concepts of the ontology. The 4 phases are detailed below.

Phase 1: Knowledge crawling from 15 websites. Brewster *et al.* [38] extracted 41 arbitrary texts from the internet on a number of artists during a task in which they evaluated an ontology built to represent a domain of artists and artefacts [7]. A similar approach is used for evaluating our barriers ontology (the conceptual barrier model). We automatically extract domain specific information by crawling 15 websites obtained through the Google search engine⁵ via the search query '*barriers to physical activity for T2D*' using the Google search library in python⁶. Important to note that, these 15 different sources (Appendix B.2) are completely exclusive of the original set of 19 sources used in the systematic review (Section 4.3.2). The information extracted is in Hypertext Markup Language (HTML)

²<https://pubmed.ncbi.nlm.nih.gov/>

³<https://bjgp.org/content/60/577/570>

⁴<https://clinicaltrials.gov/ct2/show/NCT01701570>

⁵<https://www.google.com/>

⁶<https://stackoverflow.com/questions/38635419/searching-in-google-with-python>

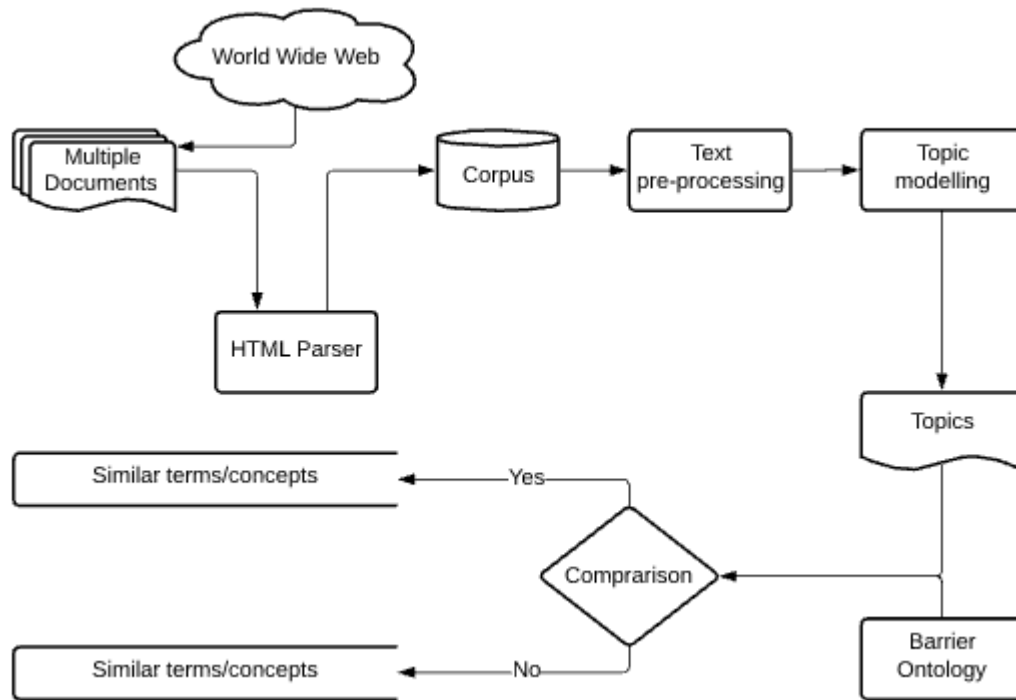


Figure 5.1: Process diagram detailing the application of the data-driven approach in evaluating the Barrier Ontology.

format, with only the body section retained. Additionally, image tags and their content are dropped, and data in table rows are extracted by accessing row tags. All the plain text extracted from each web page is merged into a single corpus of text.

Phase 2: Processing and cleaning up the corpus. The corpus obtained from phase 1 is subjected to a series of pre-processing tasks, in order to prepare it for analysis and evaluation. The state-of-the-art text mining libraries spaCY⁷ and Natural Language Toolkit (NLTK)⁸ are used, which are freely available in Python [5]. The text is then tokenized i.e. sentences are split into individual words (tokens), and each word is assigned a Part of Speech (POS) tag⁹. Assigning POS tags allows tracking of words associated to parts of speech that are not relevant to our evaluation, and henceforth the elimination of them. Categories of

⁷<https://spacy.io/>

⁸<https://www.nltk.org/>

⁹<https://spacy.io/api/annotation>

eliminated elements include conjunctions, punctuation, pronouns, and infinitive markers. The POS tags assigned to these categories are ADP, PUNCT, PRON and PART, respectively. This is followed by removal of stop words referenced from WordNet [173]. Stopwords [173] are words that don't add any semantic knowledge during analysis of the text (e.g. "an", "it", "for", etc.).

Phase 3: Topic modelling. Topic modelling is referred to as the division of text or discourse into topically coherent segments [215]. It is used in machine learning to discover semantic structures in text, hence discovering topic representations of a collection of text. As part of the analysis, several topics/keywords or themes are extracted from the cleaned corpus, which are representative of the domain, based on the sites crawled. Keyword extraction is achieved using the Latent Dirichlet Allocation (LDA) [33] probabilistic model, with the help of text mining library Gensim¹⁰. It's noteworthy to mention that we filter out extracted topics/terms that are irrelevant to the domains of focus (barriers to physical activity). Frequent examples of terms extracted include 'hypoglycaemia' or 'blood glucose levels', 'heart disease' and 'obesity', whereas irrelevant terms filtered out include 'thematic analysis', 'saturation' and others.

Phase 4: Identifying overlapping terms. Brewster *et al.* [38] used the *precision* and *recall* metrics to measure the overall fit of an ontology to the corpus. In Information Retrieval (IR) research, *precision* is the proportion of retrieved documents that are relevant, while *recall* is the proportion of relevant documents that are retrieved [258].

In addition to these two metrics, we also compute the *F1-score*, which is defined as the harmonic mean of precision and recall, and is known to be more informative in IR experiments (computed by equation 5.3) [277]. We therefore use these three measures of lexical keyword coverage of the Barrier Ontology concepts on the corpus representing the domain. Identifying overlapping terms also requires the annotation of similar concepts and differing concepts across the two representations (ontology and corpus).

$$Precision = \frac{TruePositives}{TruePositives + FalsePositive} \quad (5.1)$$

$$Recall = \frac{TruePositives}{TruePositives + FalseNegatives} \quad (5.2)$$

¹⁰<https://radimrehurek.com/gensim/>

$$F1\text{-score} = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (5.3)$$

The variable True Positives (TP) indicates the number of relevant terms that are correctly retrieved; False Positives (FP) indicates the number of incorrectly retrieved terms i.e. terms that were retrieved, but not relevant so should have not been retrieved; False Negatives (FN) indicates the number of relevant terms that were not retrieved, but should have been retrieved. The selection criteria for similarities between the Barrier Ontology and the corpus is as follows.

Syntactical similarity via string similarity. String similarity is often used as a means of quantifying the likelihood that two pairs of strings have the same underlying meaning, based purely on the character composition of the two strings [26]. Most string similarity techniques involve matching strings character by character. For instance, given an aligned pair of strings ‘barrier’ and ‘barriers’, string similarity declares that the two terms differ from each other by only a single character i.e. ‘s’ appearing at the end. The order of the characters in the terms is preserved during the matching process.

We adopt the Levenshtein Distance Algorithm [223] to compute the similarity between concepts extracted from the corpus and those in our ontology. Levenshtein Distance (LD) refers to the minimum number of changes (character insertion, substitution, or deletion) required to transform one string into another [223]. We use an online resource¹¹ which computes the LD as a percentage i.e. for a pair of strings x and y :

$$LD_{x,y} = (1 - \frac{l}{m}) \times 100 \quad (5.4)$$

Where l is the standard LD, and m is size of the longer word of the pair. For example, given strings $A = barrier$ and $B = barriers$, $l(A, B)$ is ‘1’, since one operation is required to transform A into B , namely the deletion of s at the final position of b . As $|B| > |A|$, $LDA, B = (1 - \frac{1}{8}) \times 100 = 87.5$. This means there is approximately 88% similarity between strings A and B . We retain only string pairs where both strings are identical, with a sample included in Table 5.1. A complete list of these syntactical similarities is provided in Appendix B, (Table B.1).

¹¹<https://asecuritysite.com/forensics/simstring>

Synonymous and Hypernymous phrases. Brewster *et al.* [38] used WordNet [173] to obtain hypernyms for ontology concepts and terms (during the query expansion step in their evaluation architecture), exhaustively comparing ontology concepts to corpus terms, yielding a measure of coverage of the ontology on the corpus. Similar to Brewster *et al.*, we use the WordNet knowledge base in NLTK [145] to identify concepts in our ontology that were either synonymous or hyponymous with concepts extracted from the corpus. WordNet is a popular large lexical database that groups words into a set of synonyms. For example, WordNet shows ‘feeling lazy’ is a synonym of ‘laziness’. We search for the synonyms and hypernyms for each of the ontology concept words and match these with terms extracted from the corpus. Terms in the corpus which both appeared in the cluster of synonyms and were shown as hypernyms (from WordNet) of the ontology concepts, were marked for similarity (Algorithm 1 reveals the mapping process used in identifying the related terms). No hyponymous phrases were identified while traversing WordNet, and as such we only report synonyms. We present a sample of these synonyms in Table 5.1. A complete list of synonymous similarities is provided in Appendix B, (Table B.1).

Algorithm 1 Search Wordnet to determine relatedness between ontology concepts and corpus terms.

Input: ontology_concepts and corpus_concepts

Output: synonymous and hypernymous terms

for each $concept_o$ in enumerate(ontology_concepts) do

 for each $term_c$ in enumerate(corpus_concepts) do

$concept_o_cluster = wordnet.synsets(concept_o) + wordnet.hypernyms(concept_o)$;

 if $term_c$ in $concept_o_cluster$

 store similar terms ($concept_o, term_c$) **end**

Table 5.1: Some similarity between terms of the Barrier Ontology and the corpus

Type of similarity	No.	Ontology	Corpus
Syntactical	1	Hypoglycemia	Hypoglycemia
	2	Heart disease	Heart disease
	3	Obesity	Obesity
Synonymous	4	Dislike	Disliking
	5	Feeling lazy	Laziness
	6	Bad weather condition	Climatic condition

Subsequently, we examine non-overlapping terms to complete the ontology-corpus

mapping process. We annotate terms in the ontology that didn't exist in the corpus, as well as those terms that were in the corpus but absent in the ontology. We additionally reviewed corpus terms absent in the ontology for domain relevance i.e. assessing whether these terms were relevant or irrelevant to the barrier domain. The irrelevant non-overlapping terms in the corpus were consequently excluded from the analysis, retaining a precise list of non-overlapping terms. Out of all the remaining domain relevant topics extracted, '19' terms (Tables 5.1 and B.1) are discovered in both the ontology and the corpus i.e. they are referred to as TP, and '0' relevant terms are found missing in the ontology, referred to as FN. '5' terms ('transportation problem', 'lack of knowledge', 'financial problem', 'preference for other activity' and 'feeling depressed') are discovered in the ontology which are absent in the corpus; these are referred to as FP. These measures are then used to obtain baseline metrics that evaluate the ontology as shown below:

$$Precision = \frac{TP}{TP + FP} = \frac{24}{24 + 19} = 0.558 (\approx 56\%)$$

$$Recall = \frac{TP}{TP + FN} = \frac{24}{24 + 0} = 1 (100\%)$$

$$F1-score = \frac{2 \times Precision \times Recall}{Precision + Recall} = \frac{2 \times 0.558 \times 1}{0.558 + 1} = 0.716 (\approx 72\%)$$

Discussion

Brewster *et al.* penalised an ontology during evaluation for disagreements between the ontology and the corpus [38]. In our scenario, we utilise three metrics as weighted measures of the agreement between the ontology and the corpus. As observed, we obtain a low precision of approximately 56%, implying that the number of terms that existed in the ontology were more than those in the corpus. Despite this, we obtained 100% recall, which implies there were no relevant terms in the corpus that didn't exist in the ontology. We noted that the F1-score is more informative than the precision and recall metrics, giving a more balanced measure of the extent of agreement between the ontology and the corpus. [96]. A low F1-score (towards 0%) implies a low vote of confidence in the ontology whereas a high F1-score (towards 100%) implies a high vote of confidence in the ontology. It's noteworthy to conclude that with an F1-score of approximately 72%, we are confident that the ontology is a good fit for the domain "barriers to physical activity for T2D patient".

Recording an F1-score nearing 100% would imply that the ontology has every term in the corpus. This is almost impossible with the methods we used in populating the ontology; the classification of the terms (topic annotation) was manual in the systematic review, whereas the concept extraction from the corpus was a semi-automatic process that relied on probabilistic modelling (LDA) to assign a topic to a keyword.

5.3 Competency Questions Technique

Competency questions (CQs) are formally defined as a set of questions expressed in natural language, which the finished ontology must be able to answer correctly [30, 99, 193]. In addition to being used to determine the domain and scope of the Barrier Ontology, as mentioned in Section 4.3, they can also be used to evaluate an ontology [261].

The CQs-based evaluation focuses on the verification of the Barrier Ontology by comparing it against its ontology specification document (ontology requirements [228], thus evaluating the ability of the barriers ontology to satisfy its functional requirements (detailed in Section 4.2). Contrary to data-driven evaluation, CQs provide a holistic assessment of the ontology as an artefact built to satisfy a given aim.

In our CQ-based evaluation, we construct four unique templates, to hold four different groups of CQs each. The CQs are transformed from natural language into machine processable queries expressed in SPARQL (Section B.3), a language for querying Resource Description Framework (RDF) datasets [138] (Section 3.5.3), that is a World Wide Web Consortium (W3C)¹² [214] recommendation.

Gomez [91] highlights that queries in information retrieval can either contain relations, or concepts, or both; the target concepts and relations we aim to evaluate are embedded in a number of SPARQL queries that translate, in a machine processable format, the CQs formulated in natural language. Often CQs cannot be directly translated into SPARQL, given that queries are expressed as sequences of triples of the type (*subject, predicate, object*), that define patterns to be matched in the populated ontology. We then construct the queries as illustrated by the following example. Considering the CQ, “What are the types of barrier that *prevent* or limit people from performing ‘*football*’, despite being interested in this activity?” we perform the following steps to create the associated SPARQL query:

1. Identify the main concepts in the CQ. In the above example, these include a person

¹²<https://www.w3.org/TR/rdf-sparql-query/>

who is a user/patient, and “football”.

2. Identify the main properties in the CQ which the main concept is relating to. In the example above, the property is the *prevents* relation.
3. Write a SPARQL query to select instances that contain the concepts in (1) and the properties in (2). Included in the template discussions below, are formally written SPARQL queries used in retrieving answers for CQs of each template.

The CQ-based evaluation conducted is validated by eliciting knowledge (in the form of answers to the CQs) from the feedback given in a survey aimed at experts in the health domain such as general practice doctors, clinical and medical academics or researchers, and other T2D knowledgeable people.

The domain expert knowledge is used in comprehensively validating the answers (generated by executing the SPARQL queries) to the CQs, following [30, 193], i.e. we evaluate the extent to which answers to the CQs modelled in the ontology are representative of the views of the domain experts.

The answers to the SPARQL query identify essentially either the domain or the range of the CQ’s candidate relation. We firstly evaluate the accuracy of the domain or range depending on which of the two was the appropriate answer to the question. Secondly, we determine how comparable these answers (domain or range) are to the expected pre-defined answers (from the experts’ survey responses) to the CQ in question. For example, a CQ “what barriers prevent or limit people from performing ‘Football’, despite being interested in this activity?”, needs to query the ontology and retrieve the domain of the *prevents* relation given the range ‘*football*’. The evaluation verifies the accuracy or precision of the relevant domain as well as how it correlates with the domain expert answers for the same question.

Evaluating the precision of the domain enables us validate the semantic coherence of the relational properties in the ontology, by ensuring that the meaning of the relation satisfies the constraints that are identified by experts in the survey. For example, in the case of the *prevents* relation, barriers that prevent athletics sports (e.g. football) might not necessarily prevent conditioning exercises (e.g. swimming).

Section 5.3.1 discusses the CQs and survey questions, where the latter are directly derived from the former. This relationship between CQs and survey question is presented in Table 5.2. In addition to the survey aims, Section 5.3.1 also presents the four CQ templates used in the evaluation to model the different categories of CQs, with each template accompanied by a discussion assessing the relation between the answers gathered

from the survey, and those retrieved from the ontology for the particular CQs. A detailed description of the survey design and preparation is presented in Appendix B.3.1.

5.3.1 Survey

Qualitative evaluation is widely adopted in both scientific and non-scientific fields of research. Online surveys are one of the many qualitative evaluation tools often employed by ontology engineers in evaluating the fitness for purpose of an ontology [129]. As a part of the evaluation in this thesis, we obtain institutional ethical approval from the University of Liverpool Department of Ethics¹³ to conduct an online survey titled “Ontology to Support Behaviour Change Intervention: Barriers to Activity”.

The survey questionnaire is a mixture of both structured (fixed/multiple-choice response) and non-structured (open-ended) questions. The survey is designed using a web-based platform (Allcounted.com¹⁴). Further discussion of the survey design and content are included in Appendix B.3.1. The survey invites participants (targeted through emails and social media platforms) who are oriented with medical-practice and medical-research, and T2D knowledge (further details on the participants demographics is presented below). The survey’s aims and results are presented and analysed as a crucial stage of the CQ-based evaluation. Tables 5.2 and B.3 demonstrate the relation between functional requirements, template CQs, survey questions, and target relations. The template CQs inherently reflect particular functional requirements (drawn from the ontology specification document). For presentation purposes, we only present a subset of the CQs and present the rest in appendix B.3.2. The survey questions are derived from the template CQs, and finally the target relation is the candidate relation (inserted in the query of a CQ to the ontology) subject to evaluation for the CQs of template in question. The meta-data elaborates the relationship between the CQs and the survey question.

Survey Aims

The survey was designed to elicit knowledge from domain experts, which would be used in objectively validating the scope and authenticity of the functional requirements of the Barrier Ontology. Analysis of the survey feedback is chiefly used to:

1. Assess the correlation between the domain expert knowledge presented in the survey

¹³<https://www.liverpool.ac.uk/research-integrity/research-ethics/>

¹⁴<https://www.allcounted.com/>

Table 5.2: A subset of each template’s CQs, with their corresponding survey questions, target relations, and the ontological functional requirements they satisfy. Additional comments narrating how survey questions are derived from CQs are included as Meta-data. The complete table covering all functional requirements and survey questions is in Appendix (Table B.3).

Functional require- ment	To recognise barriers that <i>prevent</i> performance of a specific type of activity.
Templates 3	CQ 3.1: What are the barriers that prevent or limit people from performing ‘ <i>Football</i> ’, despite being interested in this activity? CQ 3.2: What are the barriers that prevent or limit people from performing ‘ <i>Swimming</i> ’, despite being interested in this activity?
Survey question	For each barrier (listed on the left), please select which activity types (listed along the top) you believe the barrier could <i>prevent</i> people to engage into. For each barrier, you may select multiple answers. Leave blank the boxes for any barrier for which you are unsure.
Meta-data	The 2 CQs require barriers that prevent a single physical activity each i.e. Football (CQ 3.1), Swimming (CQ 4.2) etc. However, we design the corresponding survey question to probe for barriers that categorically prevent several physical activities presented in CQs within the template, i.e. the survey question requires barriers that prevent aerobic athletic sports (e.g. Football for CQ 3.1) and conditioning exercises (e.g. Swimming for CQ 3.2) etc.
Target relation	<i>prevents</i> relation

feedback and the answers that the Barrier Ontology provides to the CQs. For example, given a CQ that probes for physical activities suitable for barriers for Type 2 Diabetes (T2D) patents, we aim to verify that the Barrier Ontology (our model) provides suggestions of physical activities that have a bearing to those recommended by experts with clinical practice or medical research knowledge.

2. Leverage on the comparison of domain expert knowledge with ontology provided results, in order to evaluate the accuracy of domain and range as well as the semantic coherence of the relations that characterise the CQs. This process investigates whether querying the relations (that characterise the CQs) retrieves answers from the ontology

as expected by the experts in the survey.

3. Finally, the survey evaluation aims to validate the employed CQs as suitable and relevant for the Knowledge-Based System (KBS) which underpins the studied domain (Barrier Ontology domain).

In addition, the survey results could be used to reconcile any contradictory knowledge and limitation of the systematic review, particularly with regard to the barrier concept where classifications for some identified barriers, such as ‘body image’ and ‘obesity’, were not found within the reviewed studies (Section 4.3.2).

Demographics of the survey participants

A total of 30 participants completed the survey. These participants include many different medical professions:

- *Health psychologists*, who are members of the health care team who have much to contribute to the well-being and welfare of people [221];
- *Clinical psychologists*, who are mental health professionals with highly specialized training in the diagnosis and psychological treatment of mental, behavioural and emotional illnesses [175];
- *Psychiatrists*, who are trained doctors who have specialised in the field of diagnosing and managing mental illnesses, mental disorders and emotional and behavioural disturbances [31];
- *Endocrinologists* or other diabetes specialists, who are medical doctors that specialize in the glands of the endocrine (hormone) system [263];
- *General medical doctors*¹⁵ such as GPs who treat patients with minor and chronic illnesses or disease;
- Academic researchers in fields such as behaviour and behaviour change;
- *Allied health professionals*¹⁶ such as physical therapists, occupational therapists and respiratory therapists;
- *Biomedical scientists*¹⁷ who work specifically in the context of medicine.

As Figure 5.2 illustrates, general medical Doctors accounted for 24.14% of participants, as did academic researchers, making the joint-largest groups of participants. The second

¹⁵<https://www.england.nhs.uk/gp/>

¹⁶<https://www.england.nhs.uk/ahp/role/>

¹⁷<https://www.healthcareers.nhs.uk/explore-roles/healthcare-science/roles-healthcare-science/life-sciences/biomedical-science>

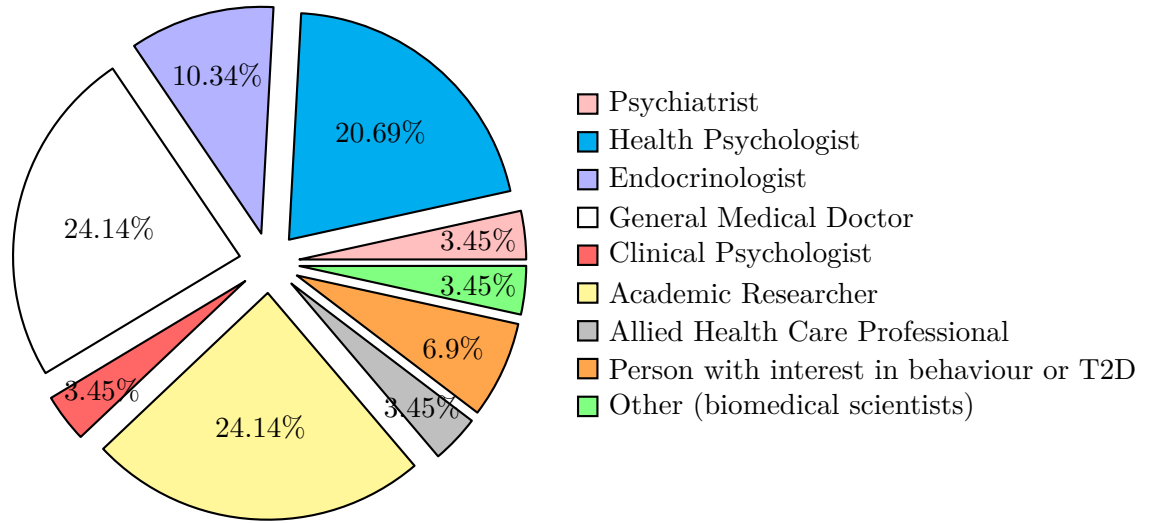


Figure 5.2: Participation per profession distribution for the “Ontology to Support Behaviour Change Intervention: Barriers to Activity”.

largest group was health psychologists, accounting for 20.69%. Smaller groups included clinical psychologists, psychiatrists, allied health professionals and biomedical scientists, each accounting for 3.45%. Other participants were split between endocrinologists and interested persons with T2D knowledge, as shown in Figure 5.2. Among the 30 participants, 1 participant partially completed the survey by leaving 1 question unanswered (scenario 3 as presented in Appendix B.3.2).

5.3.2 Analysis of Survey Results for Template Competency Question

Determining a sample size for analysis of e-survey results is not a cut-and-dry procedure, however the larger the sample, the more likely it is to yield statistically significant results [110]. According to [110, 149], sample size of at least 30 is sufficient.

In order to decide on the response that is to be recorded as the given survey results answer, we adopt a majority vote mechanism which selects the answer with the highest number of nominations for each question. An example of this strategy’s implication is: if 80% of the participants classify ‘pain’ as a ‘physical barrier’ and 20% classify it as a ‘health barrier’, we report ‘pain’ as a ‘physical barrier’. Isaac and Michael [116] recommend sample sizes ranging from 10-30 as sufficient in testing a null hypothesis of a small sample

survey. We therefore set the minimum threshold of votes (survey responses) required to consider an answer at 10 votes (approximately 33% of 30 participants), meaning answers that have less than 10 votes were excluded from the analysis.

Having received responses from a total of 30 participants, an analysis comparing answers to CQs retrieved from the Barrier Ontology to the answers gathered from the domain experts in the survey is performed. We determine a similarity score as the number of exact matching items within answers retrieved from the two sources (ontology and survey). In order to summarise this aggregated information in our work, we define Percentage Similarity (PS) as the similarity score expressed as a percentage of the total number of unique retrieved items from the two sources. The PS value serves as a functional measure of the consistency, or concurrence, between the Barrier Ontology retrieved answers and the survey gathered answers, with respect to CQs.

The following sections give a comparison of the Barrier Ontology retrieved results to Survey retrieved results to CQs, categorised under four templates. Acronyms BO and SR are used to refer to the Barrier Ontology results and Survey Results respectively, and PS refers to the Percentage Similarity score between BO and SR. Four CQs are prepared for each template, totalling 16 CQs, however for each template we present comparison between the results of BO and SR of only 2 CQs. The remaining 2 CQs under each template and the SPARQL queries of all the 16 CQs are presented in Appendix B (Section B.3). Table 5.2 shows a subset of each template, with their corresponding survey questions, with the complete table being presented in Appendix B (Table B.3). Comprehensive survey results corresponding to the CQs are presented in Appendix B (Section B.3.2).

Criteria for Evaluation. Given a CQ to which we have obtained two sets of answers, (the first retrieved from the Barrier Ontology (BO) and the second from the survey (SR)), we report the degree to which BO matches SR as a percentage similarity (PS) value introduced in preceding paragraph. A lower PS value indicates that the knowledge base (BO) does not effectively represent ground truth domain knowledge from the domain experts. A higher PS value indicates that the following assertions are more likely to be valid: (1) the relational properties (relations) are a good representation of the relationships between the concepts in question with respect to the domain expert feedback (SR), as it pertains to the scope or domain that the ontology is built for; (2) the semantic coherence of the relations in the ontology; and (3) the ontology is correctly addressing the CQ and hence able to meet its intended purpose.

5.3.2.1 Template 1

Template 1 CQs precisely query the ontology for barriers that belong to a single barrier type or category. Two example CQs under this template are:

CQ 1.1: What barriers are classified as Health barrier?

CQ 1.2: What barriers are classified as Environmental barrier?

Table 5.3 reveals answers corresponding to CQ 1.1 and CQ 1.2, in terms of classification choices made by the experts (SR) along with those retrieved from BO, including the percentage similarity (PS).

Table 5.3: Classifications assigned to barriers according to the Barrier Ontology (BO) results and Survey Result (SR) for CQs 1.1 and 1.2 under Template 1. ✓implies a barrier is identified as belonging to the corresponding Barrier Type, whereas X is the opposite. A match (Percentage Similarity (PS)) between the BO and SR column indicates that the ontology retrieved answers identical to those provided by experts in the survey.

CQ No.	Barrier Type	Barriers	BO	SR	PS %
CQ 1.1	Health Barrier	Hypoglycaemia	✓	✓	40%
		Medical condition	✓	✓	
		Fatigue	X	✓	
		Pain	X	✓	
		Obesity	X	✓	
CQ 1.2	Environmental Barrier	Transportation difficulty	✓	✓	100%
		Lack of facility/equipment	✓	✓	
		Facilities/equipment cost	✓	✓	
		Lack of safety	✓	✓	
		Bad weather condition	✓	✓	

Discussion. Given CQ 1.1 (What barriers are classified as ‘*Health*’ barrier?), we notice that while survey results classified 5 barriers as ‘*health*’ barrier, the Barrier Ontology only classified 2 barriers as ‘*health*’ barrier, resulting in a PS of 40% ($\frac{2}{5} \times 100 = 40\%$) as shown in Table 5.3. On the other hand, a PS of 100% is obtained for CQ 1.2 since the results from BO and SR were identical.

As shown in Figure 5.3, there was a high PS between the two reviewed samples (BO results and SR results) for CQ 1.2 (which probed for environment barrier). CQ 1.1 (health barrier) and CQ 1.4 (physical barrier) (whose results are presented in Appendix B.3.2) were however noticeable outliers, with a PS value of 40% and 67%, respectively. The low

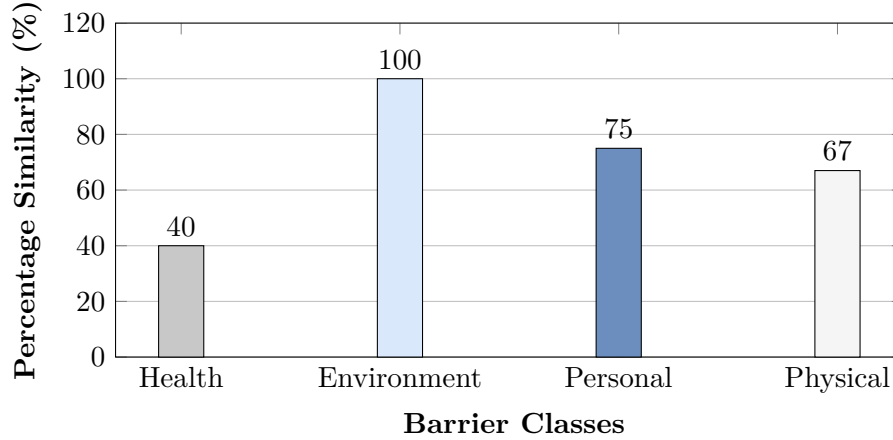


Figure 5.3: Bar chart depicting the agreement rate between the retrievable BO results and the SR for the CQs classified under Template 1.

PS score of Template 1 CQs was attributed to two main reasons: (1) some barriers did not exist in the systematic review and therefore unclassified in the ontology, whereas the same were assigned categories in SR. For example, ‘obesity’ and ‘body image’ lacked a category to which they belonged in the ontology, whereas SR classified them as health and psychological barriers respectively; (2) while BO strictly categorizes each barrier into a single barrier type, SR sometimes classified barriers into multiple barrier types e.g. barriers ‘dislike’ and ‘laziness’ are classified as both ‘personal barriers’ and ‘psychological barriers’. In this case, BO results were found to only match one of the two barrier types determined by SR. These two limitations (unclassified barrier and inconsistent barrier classifications) are discussed in Chapter 6 (Section 6.4).

CQs in Template 1 are characterised by the *subClassOf* relation (i.e. the target property identified in the competency question and used in formulation of the query that retrieves answers from the ontology). In conclusion, we indicate that the *subClassOf* relation appropriately captures the superior-subordinate relationships that exist between the environment, personal barrier categories and their respective sub-barriers, with respect to the domain expert knowledge. However, for the reasons of the low PS score stated above, the relation does not capture all the sub-barriers under the health, physical and barrier categories.

5.3.2.2 Template 2

Template 2 constitutes CQs which query the ontology for barriers that prevent a patient from performing a physical activity, based on the patient's characteristics. Two example CQs under this template are:

CQ 2.1: What is the expected barrier for a male patient who is 37 years old, has a full-time job, lives with his wife and two children, and complains about not having friends to play sports with, and no personal support for performing physical activity?

CQ 2.2: (What is the expected barrier for a person who has a full-time job, a minor injury, and lives in a very crowded area with traffic issues? Their nearest sports centre or gym and walking trails are more than an hour away by bus?).

Table 5.4 reveals answers corresponding to the above CQs from SR, along with BO retrieved results.

Table 5.4: Potential barriers according to the Barrier Ontology (BO) and Survey Results (SR) for scenarios 1 and 2 under Template 3. ✓implies a barrier is a hindrance to patient in response to the CQs, whereas X is the opposite.

Barriers	Scenario 1 (CQ 2.1)		Scenario 2 (CQ 2.2)	
	BO	SR	BO	SR
Lack of time	✓	✓	✓	✓
Lack of support	✓	✓	-	-
Lack of facility	-	-	✓	✓
Trans. problem	-	-	✓	✓
Pain e.g. injury	-	-	✓	X
Percentage Similarity (PS)	100 %		75 %	

Discussion. With reference to Table 5.4 and Figure 5.4, It is observed that answers from the BO and SR are identical for CQ 2.1, producing a PS of 100%, whereas for CQ 2.2, only 3 out of 4 answers were found to be identical between answers from BO and SR, producing a PS of 75%.

CQs in Template 2 are characterised by DataType Properties (DTP) (i.e. the target properties identified in the CQ and used in formulation of the query that retrieves answers from the ontology). For example in CQ 2.1, the identified DTPs in the BO taxonomic structure include: age, isMarriedWithChildren, hasFullTimeJob and isSocial. These DTPs are respectively assigned values 'under 60', 'yes', 'yes' and 'no' respectively in the query used to retrieve the barriers that the patient suffers. Upon comparing the BO retrieved

barriers (i.e. barriers linked to these DTP values) to barriers in SR, a maximum PS score is obtained indicating that the relations (DTPs) and their literals are good representations of the domain expert knowledge. On the other hand, for CQs 2.2 and 2.4 (Figure 5.4) a lower PS value was obtained, signalling that knowledge which encapsulates the DTPs responsible for answering the CQs 2.2 and 2.4 was inconsistent with the domain expert knowledge. CQs in this template were characterised by 2 or more data properties reflecting the multiple patient characteristics quoted in the CQs, that each contribute to the eventual answer. We attribute the lower PS value to the aforementioned ambiguity and complex nature of the CQ sentences for this template.

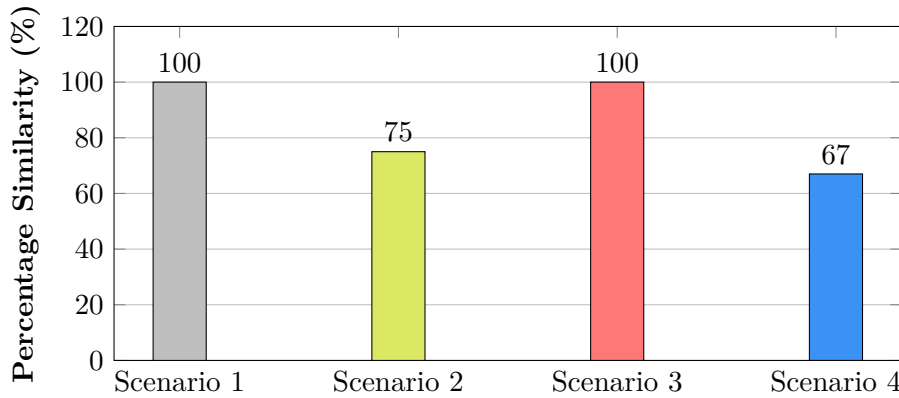


Figure 5.4: Bar chart depicting the percentage similarity between the retrievable BO results and the SR for the CQs under Template 3.

5.3.2.3 Template 3

Template 3 consists of CQs which query for barriers that prevent a patient from performing a specific physical activity. Two example CQs for this template are:

CQ 3.1: What barriers prevent or limit people from performing ‘Football’, despite being interested in this activity?

Question 3.2: What barriers prevent or limit people from performing ‘Swimming’, despite being interested in this activity?

Tables 5.5 and 5.6 show the answers from both BO and SR for CQs 3.1 and 3.2, including the percentage similarity.

Discussion. Tables 5.5 and 5.6 reveal a high overlap in the selections made by the experts (SR) as well as answers retrieved from the ontology (BO) to CQ 3.1 (Football) and 3.2 (Swimming), producing a PS of 83.33% and 82.60% respectively. CQs under Template 3 are characterised by the *prevents* relation (the target property identified in the CQ), i.e. they probe for the triples ‘Barrier *prevents* Physical Activity’, particularly retrieving the domain of the *prevents* relation (barrier) given the range (physical activity).

The high PS scores (of at least 82%) as demonstrated in Figure 5.5 signalled that triples of the *prevents* relation, particularly ‘Barrier *prevents* PhysicalActivity (where $\{Football, Swimming, Yardwork, Running\} \in PhysicalActivity$)’, reflect to a good degree the knowledge of the domain experts in terms of which particular activities prevent which particular barriers. The marginal differences between the BO and SR across all 4 CQs in this template were attributed to the following: (1) knowledge representation is subjective, especially in specialised domains such as health i.e. it can vary from one developer to another. In our case, some classifications and relations inferred from the systematic review (hence BO) differed slightly from those inferred from the SR; (2) Preferably, incorporating domain expert knowledge at the ontology development phase would have ideally enhanced the process of drawing knowledge pertaining to activity preventing barriers, which would increase the chances of having higher PS during evaluation. This means that the knowledge base on which the Barrier Ontology is built is now found to be less comprehensive than required, though is still indeed sufficient (due to the good PS values that we obtain).

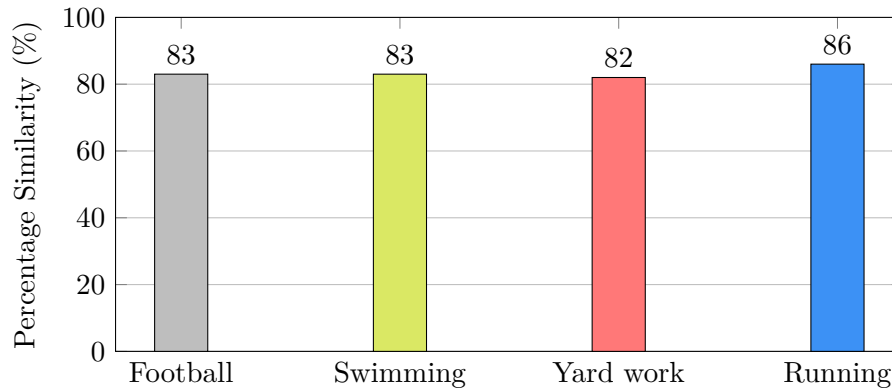


Figure 5.5: Bar chart depicting the agreement rate between the retrievable BO results and the survey results for all CQs under Template 3.

5.3.2.4 Template 4

Template 4 CQs query the ontology for activities suitable for patients who encounter specific barriers. Two example CQs under this template are:

CQ 4.1: What physical activity is suggested for the bad weather condition ('Raining') barrier?

CQ 4.2: What physical activity is suggested for the lack of time ('Home responsibilities') barrier?

Table 5.7 compares suggested alternative activities, including the percentage similarity, for the *bad weather condition (Raining)* barrier (CQ 4.1) and *Financial problem* barrier (CQ 4.2) for BO and SR.

Table 5.7: Suggested activities according to the Barrier Ontology (BO) and Survey Results (SR) to limit bad weather condition barrier (CQ 4.1) and lack of time barrier (CQ 4.2) under Template 4. ✓ implies an activity is a suggestion to patient in response to the CQs, whereas X is the opposite.

	Bad weather (CQ 4.1)		Lack of time (CQ 4.2)	
Physical activities	BO	SR	BO	SR
Aerobic exercises	-	-	✓	✓
Occupational activities	-	-	X	✓
Conditioning exercises	✓	✓	-	-
Home activities	✓	✓	✓	✓
Percentage Similarity (PS)	100%		67%	

Discussion. Table 5.7 reveals a maximum PS of 100% for CQ 4.1, showing that both BO and SR suggested the same activities for patients who encounter the Bad weather barrier. On the other hand, the table also reveals a lower PS score of 67% for CQ 4.2, as a result of BO and SR having only 2 identical barrier selections out of 3. Template 4 CQs are characterised by the *isSuggestedFor* relation (the target property identified in the CQ), i.e. they query for triples ‘Physical Activity *isSuggestedFor* Barrier’ particularly retrieving the domain of this *isSuggestedFor* relation (Physical Activity) given the range (Barrier). For example, CQ 4.1 “What physical activity is suggested for the bad weather (Raining) barrier?” queries for the domain of the *isSuggestedFor* relation where the range is “badWeatherCondition”. The significant PS scores obtained for CQ 4.1 (Bad weather) and CQ 4.3 (Financial problem) (as demonstrated in Figure 5.6), indicated that the triples of

the *isSuggestedFor* relation modelled in the ontology are reflective of the domain expert knowledge for these particular CQs, whereas the low PS scores obtained in CQ 4.2 and CQ 4.4 (as also demonstrated in Figure 5.6) indicated the knowledge captured by the relation was inconsistent with the domain expert knowledge for these CQs. Similar to the inconsistency between BO and SR in Template 3, this inconsistency was also attributed to the fact that knowledge representation is subjective, especially in specialised domains such as health, and that knowledge pertaining to suitable activities for barriers in the BO (inferred from the systematic review) varied from that captured from the experts in SR.

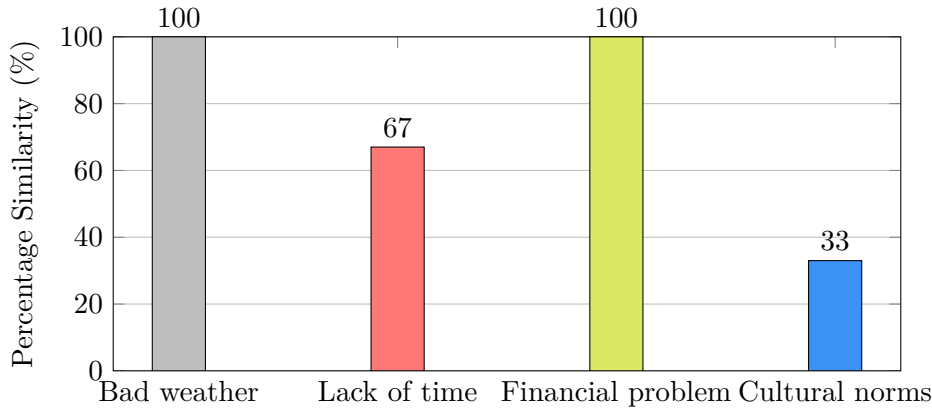


Figure 5.6: Bar chart depicting the agreement rate between the retrievable BO results and the survey results for all CQs under Template 4.

5.3.3 Conclusion of Competency Question Evaluation

The CQs technique evaluates the ability of the Barrier Ontology to satisfy its functional requirements (Section 5.2). A set of CQs corresponding to different functional requirements were constructed. Using SPARQL, answers to the CQs were successfully retrieved from the ontology, hence validating the functional purpose set out for the ontology, as discussed in Chapter 4 (Section 4.2). After this, we probed for the authenticity and validity of the retrievable results by eliciting domain expert knowledge (from a survey) in form of answers to the same CQs. This approach is often referred to as *assessment by humans* in previous studies [37].

We quantify the agreement between 30 domain experts and the ontology by using a Percentage Similarity (PS) value, which is the number of exact matching items within

answers retrieved from the two sources (ontology and survey) expressed as a percentage of the combined total number of unique answers. This PS metric is a measure of relatedness or agreement between domain expert answers and ontology’s answers to 16 CQs categorised under 4 different templates. As observed in the results presented, we report a significant agreement between the two for majority of the CQs. However, we also report a disagreement on two CQs (1.4 and 4.4), where a low PS was discovered between results from the ontology and those from the Survey. Outstanding similarity scores are obtained for Template 3 CQs, implying domain experts largely agree with the ontology’s content retrieved for the template CQs.

As a part of the survey feedback, the survey was notably commended by the domain experts, with many of them expressing their satisfaction in the overall questioning.

5.4 Summary

This chapter embarked on validating the developed Barrier Ontology as a classification structure representative of the domain *‘barriers to physical activity for T2D patients’*. The chapter extensively narrates how two methodologies are used for evaluation of the Barrier Ontology, namely data-driven evaluation and competency question-based evaluation. Using the data-driven strategy, we evaluated the congruence of the Barrier Ontology with a corpus (collection of text) that was representative of the domain we investigate. This corpus was generated from 15 website articles that were completely exclusive of the articles we used in the systematic review (Section 4.3.2). We report a high score of the harmonic mean between precision and recall (72%) which is emblematic of the co-occurrence of terms across both the ontology and the corpus, thus giving us confidence that the ontology successfully conceptualised the subject domain as required by the first sub research question of the thesis (How can we build the Barrier Ontology to specifically conceptualise the subject domain *“barriers to physical activity for T2D?”*).

The CQs-based evaluation focused on evaluating whether the ontology achieves its purpose as declared in the functional requirements. We probed for answers to a set of 16 CQs that reflected the functional requirements. These answers were retrieved from the Barrier Ontology and from 30 domain experts via an online survey. The answers from the two avenues (Barrier Ontology and domain experts) were carefully compared mainly to measure their extent of agreement. We report generally high agreement, in terms of a PS value, between the two avenues for answers to the CQs. For instance, a PS of 82.60% was

obtained for CQs (e.g. CQ 3.2) that probed for barriers which prevent physical activity. Based on the encouraging results of the evaluation discussed above, it can therefore be conclusively argued that the ontology is a good fit for the domain.

Within this chapter, the fourth sub-research question (How can we evaluate our methodology, and offer general suggestions for ontology developers?) is answered by using the two techniques (data-driven and CQs) to evaluate the Barrier Ontology.

Chapter 6

Conclusion

6.1 Introduction

This thesis addressed a gap in the research on the digitalisation of behavioural health interventions, by concentrating on a model of the barriers which inhibit successful adoption of behaviour intervention techniques. We highlighted that prior work, most notably the taxonomy of Behaviour Change Techniques (BCT) [168], placed much emphasis in modelling behaviour change and maintenance, yet *barriers* to said behaviour change have received relatively little attention. This challenge chiefly motivates the work covered in the study, that is the design and delivery of a suitable knowledge base of behavioural barriers, which can then complement existing similar artefacts conceptualising behavioural models. This culminates into a novel "Barrier Ontology", built around the scenario of physical activity behaviour for Type 2 Diabetes (T2D), and evaluated by means of a number of techniques.

This chapter summarises the work presented in this thesis. Section 6.2 provides a summary of each chapter, Section 6.3 reviews the questions in the thesis and show how these were addressed and answered, followed by a discussion on the limitations encountered while undertaking the work. Finally, Section 6.5 discusses possible future directions that build upon the work presented in the thesis.

6.2 Summary of Thesis

Chapter 1 set the scene for the thesis, giving a general overview of topics related to the study, from the behavioural psychology literature, such as behaviour, behaviour change,

behaviour change techniques, and from the scenario literature, on the impact of physical activity on Type 2 Diabetes (T2D). This served as background for presenting the research problem and motivation for the study, which was summarised in the research question: “How can we use ontologies to formalise the impact barriers have on behaviour change (developed in psychology) and underlying assumptions in a machine-readable format, to support health informatics applications?”. This research question was further divided into several sub-questions, handled throughout the thesis. Chapter 1 also summarises key contributions towards the behavioural and behavioural change domain that emerge from this thesis.

Chapter 2 presented more thoroughly background material relevant to the thesis. The chapter started with a review of human behaviour and its complexity, being as it is influenced by several external and internal factors. Three seminal models to understand behaviour were described, along with a discussion on how it is possible to legally and ethically influence either change or maintenance of particular behaviour activities: the Capability, Opportunity and Motivation-Behaviour (COM-B) model, the Health Belief Model (HBM), and the Transtheoretical model (TTM) (or stage of change model). We also introduced an important collaborative effort to report on interventions to change behaviour, the Behaviour Change Techniques taxonomy [168], which constitutes the point of departure for our study, and we observed that the notion of barriers is not sufficiently represented, and how this potentially prevents interventions from being fully successful in the long term.

Chapter 3 moves to the more computational side of the work, and introduced research on Ontology engineering, starting with presenting the various definitions and terminology around computational ontologies, and then exhaustively details the “Ontology Development 101” methodology [192], which would eventually guide the development of the “Barriers Ontology”, output of our work. This chapter also discusses the evaluation techniques that can be used for evaluating ontologies. Finally, some efforts consisting of either taxonomies or ontologies related to behaviour and behaviour change were reviewed.

Chapter 4 provides a full account of the process and final result of the task of creating an ontology of barriers, starting with the functional and non-functional requirements, then introducing the high-level conceptual model, and finally detailing the development of the Barrier Ontology itself. This included also the description of the domain and scope of the ontology, which we identified barriers to the physical activity behaviour and managing T2D. The chapter describes a systematic review that was carried out, aimed to acquire the terms and relations, as well as the high level concepts of the ontology, and the existing

ontologies that were re-used, in conformity with the best practices described in the “Ontology Development 101” methodology that we have followed.

Chapter 5 provides a full evaluation of the Barrier Ontology, created in Chapter 4. Particularly, it thoroughly discusses the methodologies selected for evaluation of the Barrier Ontology i.e. data-driven evaluation and competency question (CQ) evaluation. The data-driven approach aims to align the ontology with a separate collection of text (corpus) representative of the ontology’s domain. The alignment was performed by measuring the semantic similarity and the syntactical similarity of words across the ontology and the corpus. The CQs approach involved formulating a set of questions which the ontology was expected to correctly address. We therefore evaluated the ontology by examining how comparable the answers it provides were to answers from domain experts (through an online survey questionnaire) for each CQ. The chapter concluded indicating how both evaluation approaches successfully proved the suitability of purpose of the Barrier Ontology.

6.3 Main Findings and Contributions

In this section, we consolidate the main contributions of this work. The contributions of this thesis come from research that was carried out in order to answer the following central research question:

“How can we use ontologies to formalise the notion of barriers to behaviour change and their underlying assumptions in a machine-readable format to support health informatics applications?”

The central question was divided into several sub-questions, each contributing to answering the central research question. In answering the research questions, scientific results and findings were obtained. We first list the contributions, before presenting each sub-question independently, presenting the associated findings in terms of the particular question. We then present the results in terms of the initial, central research question that drives the thesis. The main contributions of the thesis are:

1. A taxonomy of generic barriers to behavioural change, which provides the foundation for the ontology.
2. The identification of assumptions within the barrier’s domain knowledge, and their explicit modelling within the ontology.
3. A comprehensive analysis that annotates relations between barriers and physical activities for T2D patients, enabling researchers and software developers to gain a

greater understanding of the underlying information structure.

4. The introduction of a hybrid re-usable framework for building and evaluating an ontology driven from data.

‘How can we build an ontology of “barriers to physical activity for T2D patients?”. The thesis proposes a conceptual barrier model (Section 4.3) that encapsulates high level information pertaining to the domain i.e. barriers, physical activities, and T2D patients. Using a strategic systematic review, we build a vocabulary of terms used in the domain, identifying six most common barrier categories including health, physical, environmental, psychological, and personal barriers. We further identified and mapped terms within the vocabulary to the different barrier categories, forming a hierarchical classification. Using Tables A.1 and A.2, we provide evidential quotes (extracted from various references reviewed) necessary for supporting the induced ‘subClassOf’ relation between a term and the barrier category. These induced hierarchies were directly inherited during the development of the Barriers Ontology.

‘How can we demonstrate the use of a formal methodology, including the notion of ontology reuse, to objectively support the Barrier Ontology?’. As mentioned in Chapter 3, it is possible to supplement an ontology by integration of existing ontologies, which enhances the applicability and usability of the original ontology, and eliminates the need of creating a new model from scratch. The use of this capability represents good practice in ontology development and can be quite powerful when used properly, but will also save effort [113, 192]. Therefore, existing ontologies were imported and integrated into three concepts of the conceptual model. We import the physical activity ontology, GUMO and the human disease ontology (Section 4.3.3), which present the concepts of physical activity, patient or user, and diseases, respectively.

‘How can we use the Barrier Ontology to produce suitable recommendations of physical activities which take into account barriers to such physical activity from T2D patients?’. Identifying the barriers related to behaviour change (e.g. physical activity) without managing or overcoming these barriers is not enough [163, 222]. So, in addition to identifying barriers to physical activity, this work identifies physical activities that limit the impact of these barriers. The ontology therefore addresses the requirement “of identifying barriers and suggesting physical activities to limit identified barriers”. The

systematic review is a critical building block in which we identify the relations between concepts, which in turn are necessary in satisfying the stated requirement. Evidential quotes from references supporting the relations are provided in Chapter 4. For example, ‘Indoor activities’ is a suggested activity for the ‘Cultural norms’ barrier [240]; this source also notes “exercises that can be performed easily indoors should also be introduced as an alternative to outdoor exercise”. This is one of several examples identifying relations which are later inherited into the Barrier Ontology. .

‘How can we evaluate our methodology, and offer general suggestions for ontology developers?’. Two different approaches were used to evaluate the proposed conceptual barrier model (ontology): the data-driven approach, and the CQs techniques. The results of the data-driven approach were encouraging, with 72% of the modelled barriers’ terms discovered to be similar to terms extracted from a corpus of text collected from random venues. The CQs evaluated the capability of the Barrier Ontology to achieve its purpose. We measured the similarity between results retrieved from the Barrier Ontology with those we elicited from domain experts through a survey. Our analysis revealed that in most cases, ontology-oriented results correlated with domain-expert oriented results to the CQs. Chapter 5 and Appendix B contain further details regarding this sub-question.

Returning to the central research question, we state that it was possible to formalise the barriers to behaviour change (and their assumptions) into a machine-readable format, using an ontology to support health informatics applications and the process to do so was demonstrated in the thesis.

6.4 Limitations of the Work

The deliverables produced with this thesis are evidence of how various approaches, frameworks and methodologies could be successfully used together towards a goal. Nevertheless, a number of limitations were encountered, particularly during the extensive systematic review that identified concepts and relations, as well as in the ontology evaluation phase. Below is a discussion of the main limitations, organised based on the phase of ontology development in which we encountered them:

6.4.1 Domain and Scoping Review of The Barrier Ontology

Inadequate Domain Knowledge: The high-level domain of behaviour as it pertains to health is broad and heavily resourced, however it is also very subjective and sensitive. The data gathered is sourced solely from existing literature i.e. domain experts are not consulted for any data or divergent views on the subject to enrich our data repository. Interviews with experts would have potentially broadened the scope of the collected data used in building the ontology. Furthermore, none of the project stakeholders had sufficient knowledge in health and behaviour subject matters, such as physical activity as a behaviour and Diabetes, which prolonged the process of gathering resources necessary for building the knowledge base of the Barrier Ontology.

Inconsistent Barrier Classifications: As a result of the absence of a standardized barrier classification, some of the identified barriers are classified differently by different authors. For example, ‘lack of time’ or ‘time restriction’ was classified as an environmental barrier by [74], and on other occasions it was classified as a personal barrier by [12, 25]. This limitation was managed by opting for the most popular classification for a barrier across the reviewed studies. Personal barrier was more popular than Environmental barrier for the ‘lack of time’ scenario. While this is limitation on the quality of the final output, it however testifies to the complexity of the domain, and the need of systematic approach rather than an ad-hoc solution.

Unclassified Barrier Concepts: For some of the identified barriers, none of the reviewed studies contained information that classified them into any one of the six identified barrier categories. Examples of such barriers were ‘obesity’ or ‘overweight’ and ‘body image’. Analysis of feedback from domain experts via the survey (Section 5.3.2) provided insight in the classifications suitable for the originally unclassified barrier.

6.4.2 Evaluation Phase

Ambiguity in Survey Responses: Some of the identified barriers were classified differently by different survey participants in response to the online survey question that required participants to classify barriers (covered under competency question-based evaluation). For example, the barrier “dislike” was classified as a Personal Barrier by 40% of the survey participants and as a Psychological Barrier by 57% of the participants. This inconsistency

can be dealt with by retaining the classification that received the largest amount of nominations. We note that this ambiguity mirrors the same ambiguity in the systematic survey, which again confirms this domain is a complex one.

Ambiguity in Evaluating Natural-Language Relations: Literature does not emphatically suggest any particular level of granularity to satisfactorily validate a relation as one that represents the knowledge domain of an ontology [38]. A majority of the relations are difficult to evaluate using the data driven evaluation approach because they are phrases of either 4 or more words, hence carrying a lot of contextual information; “Embarrassment is a Social Barrier”, “Facilities and equipment cost prevents Aerobic Exercise”, “Walking is suggested for Lack of time” are examples of relations in the ontology that are difficult to evaluate using a data driven approach. Probing for semantically similar relation examples to the three example sentences in a corpus was difficult to achieve, especially without a benchmark of required granularity.

6.5 Future work

In this section, we detail potential future work on the Barriers Ontology, centred around additional research questions. For building upon the Barriers Ontology, we look to answer the following questions:

1. How can we improve the maintenance techniques employed in the Barriers Ontology?
2. How can we integrate emergent artefacts and therefore expand the knowledge base of the Barrier Ontology?
3. How can we improve the readability and usability of the Barrier Ontology?
4. How can we answer the second and third sub-questions efficiently?

6.5.1 Improving Maintenance Techniques

This work successfully built an ontology to support behaviour change interventions focusing on barriers to physical activity, i.e. Barrier Ontology. This was achieved while employing the "Ontology Development 101" process, a well-known approach for systematically developing domain-based ontologies [192]. Populating the ontology was a notably laborious exercise that involved a manual systematic review of 845 studies identified from five different electronic databases. Going forward, for purposes of ontology maintenance, one should consider

adopting automatic ontology population approaches to populate the Barrier Ontology. Despite posing challenges, such as duplicated information across documents, and redundant annotations [115], we find that the automatic extraction of concepts and terms from diverse web resources is a more effective and efficient approach for determining the important terms in a domain [51, 124]. Popularly referred to as *ontology learning*, ontology developers have employed a hybrid of machine learning and statistical approaches to populate ontologies extracting concepts and relations from resources such as twitter data [118]. Currently, there is a variety of information retrieval tools that support automatically populating an ontology, which would therefore be useful in tackling our challenge. One technique we may consider is from the Artequakt project [7]. The project proposed an extraction tool that searches online documents and extracts the knowledge that matches the given classification structure, providing it in a machine-readable format to be automatically maintained in a Knowledge Base (KB) [7].

6.5.2 Expanding the Knowledge Base

Ontologies are inevitably subject to constant changes, with many researchers in the ontology engineering space claiming that ontologies are confronted by evolution [85] and [191]. In order to address this, one should plan on the addition of concepts and relations imperative to the domain that are currently absent. This essentially will involve the exploration of techniques that leverage some form of data mining, such as Natural Language Processing and Machine Learning, because of their ability to detect hidden information from large textual data [152]. An initial starting point for this particular task is utilizing readily available tools such as GRAMEXCO [85], a document classification model that performs text analysis to extract highly semantically related terms.

6.5.3 Refinement of the Barrier ontology

Ontology refinement is a vitally important maintenance strategy that can improve the readability and usability of the ontology and support its evolution to cover new unseen concepts and constraints. Having imported existing ontologies into the Barrier Ontology, i.e. the disease ontology, physical activity ontology and the general user model ontology, there's a need to refine the ontological entities in order to retain relevant elements imperative to the application domain. Refinement consists of re-learning the meaning or interpretation of the terms and concepts in the ontology, ensuring that none of the words or terms (especially

unknown words) share a similar conceptual behaviour e.g. (Hypoglycaemia and Blood glucose lowering). The Text-To-Onto ontology engineering framework [152] effectively facilitates ontology refinement, and therefore can potentially be useful in further evolution of the Barrier Ontology.

The refinement process shall further embark on critically assessing the hierarchical relations within the Barrier Ontology. Focusing on the consistency of is-a hierarchies is a recommended guideline [176]. Takeshi develops a reliable system to achieve an evaluation of is-a hierarchies (e.g. *Hypoglycaemia is-a health barrier* [75, 143]). It requires answering the following two research questions, (1) how to find components that violate the criteria consistency guidelines for classification and (2) how to formulate refinement proposals for each inconsistent component automatically.

6.5.4 Sourcing collaboration and publishing the ontology

This work undertook extensive analysis of two health subjects: physical activity as a behaviour, and a form of diabetes called T2D. This culminated into an ontology that conceptualises a domain we called barriers to physical activity for T2D patients. Therefore, upon completion, we aim to publish the stable version of the ontology. This will help to entice interested people (especially medics or health-modelling experts) to collaboratively contribute to our work, hence improve its usability. It will also help to gather feedback from research fellows who contribute to public ontology libraries or publishing platforms.

Working with domain experts would validate and cement the knowledge base we propose in this research. We plan on adopting a framework intuitively designed by Noy *et al.* [191] to support collaborative ontology editing in various modes, such as synchronous and asynchronous editing (allows users to view each other's changes). This process allows rigorous examining of the ontology's functional requirements.

Publishing the ontology to be reviewed by other researchers is a step that will be undertaken immediately. At the moment, BioPortal¹ and OBO Foundry² are the open-source repositories we are currently considering for this, however we aim to publish the Barrier Ontology to many repositories, requesting the respective communities to comment on all entities and presentation of the ontology. To our benefit, Protégé (the main development kit we adopted) has been integrated to BioPortal, thanks to Stanford Centre for Biomedical

¹<https://www.bioontology.org/>

²<http://www.obofoundry.org/>

Informatics research [190]. This subsequently has enabled bridging of the different ontology engineering lifecycle elements, such as publishing and editing. While using this platform, it's possible to populate the ontology using content from other published ontologies.

Whilst T2D is the dominant form of Diabetes, it is not the only form of the disease. As mentioned in Chapter 4, other types of diabetes include type 1 diabetes, gestational diabetes, mellitus and monogenic diabetes syndromes [16, 200]. Our future endeavours shall embark on incorporating the diabetes disease as a whole, thereby conducting case studies to analyse what limits or barriers to physical activity are prevalent among patients, irrespective of their diabetic form. The planned automatic approach of populating the ontology with new terms and concepts would be beneficial to this aim.

6.6 Summary

Motivated by the vivid fact that barriers are inadequately catered for in both recent and current behavioural models, we embarked on systematically developing a knowledge base to contain knowledge suitable for addressing behavioural barriers. We have built the “Barriers to Physical Activity for T2D Diabetes patients” ontology that we expect to complement existing behavioural change models, and ultimately enhance the interpretation attached to several theories in the domain of understanding behaviour. Information readily retrievable from the ontology includes the classifications of barriers, which barriers prevent which activities, as well as which activities limit the negative impact of barriers.

We successfully evaluated the ontology proposed using a data-driven approach and a competency questions approach, with the latter being supplemented by domain expert knowledge gathered through an online survey questionnaire. The limitations of this work were mitigated, as is clearly highlighted in Section 6.4. We expect that future work will not only maintain the ontology, but will allow for collaboration with other ontology developers and behavioural health domain experts to further the study human behavioural change.

References

- [1] Charles Abraham and Susan Michie. A taxonomy of behavior change techniques used in interventions. *Health psychology*, 27(3):379, 2008.
- [2] Barbara E Ainsworth, William L Haskell, Stephen D Herrmann, Nathanael Meckes, David R Bassett Jr, Catrine Tudor-Locke, Jennifer L Greer, Jesse Vezina, Melicia C Whitt-Glover, and Arthur S Leon. 2011 compendium of physical activities: a second update of codes and met values. *Medicine & science in sports & exercise*, 43(8):1575–1581, 2011.
- [3] Icek Ajzen et al. The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2):179–211, 1991.
- [4] Icek Ajzen, Martin Fishbein, and Robert L Heilbroner. *Understanding attitudes and predicting social behavior*, volume 278. Prentice-hall Englewood Cliffs, NJ, 1980.
- [5] Fouad Nasser A Al Omran and Christoph Treude. Choosing an nlp library for analyzing software documentation: a systematic literature review and a series of experiments. In *Proceedings of the 14th International Conference on Mining Software Repositories*, pages 187–197. IEEE Press, 2017.
- [6] Maha Al-Yahya, Remya George, and Auhood Alfaries. Ontologies in e-learning: review of the literature. *International Journal of Software Engineering and Its Applications*, 9(2):67–84, 2015.
- [7] Harith Alani, Sanghee Kim, David E Millard, Mark J Weal, Wendy Hall, Paul H Lewis, and Nigel R Shadbolt. Automatic ontology-based knowledge extraction from web documents. *IEEE Intelligent Systems*, 18(1):14–21, 2003.

-
- [8] Dolores Albarracín, Jeffrey C Gillette, Allison N Earl, Laura R Glasman, Marta R Durantini, and Moon-Ho Ho. A test of major assumptions about behavior change: a comprehensive look at the effects of passive and active hiv-prevention interventions since the beginning of the epidemic. *Psychological bulletin*, 131(6):856, 2005.
 - [9] Yousef Alfaifi, Floriana Grasso, and Valentina Tamma. Developing a motivational system to manage physical activity for type 2 diabetes. *Artificial Intelligence for Diabetes (2016)*, 22, 2016.
 - [10] Yousef Alfaifi, Floriana Grasso, and Valentina Tamma. Towards an ontology to identify barriers to physical activity for type 2 diabetes. In *Proceedings of the 2017 International Conference on Digital Health*, pages 16–20. ACM, 2017.
 - [11] Yousef Alfaifi, Floriana Grasso, and Valentina Tamma. An ontology of psychological barriers to support behaviour change. In *Proceedings of the 2018 International Conference on Digital Health*, pages 11–15. ACM, 2018.
 - [12] Habiba I Ali, Latifa M Baynouna, and Roos M Bernsen. Barriers and facilitators of weight management: perspectives of arab women at risk for type 2 diabetes. *Health & social care in the community*, 18(2):219–228, 2010.
 - [13] R. E. Allen. The concise oxford dictionary of current english. *Clarendon Press, Oxford, England*, 1997.
 - [14] Grigoris Antoniou and Frank Van Harmelen. Web ontology language: Owl. In *Handbook on ontologies*, pages 67–92. Springer, 2004.
 - [15] American Diabetes Association et al. Standards of medical care in diabetes-2006. *Diabetes care*, 29(1):S4, 2006.
 - [16] American Diabetes Association et al. Standards of medical care in diabetes—2016 abridged for primary care providers. *Clinical diabetes: a publication of the American Diabetes Association*, 34(1):3, 2016.
 - [17] Leah Avery, Darren Flynn, Stephan U Dombrowski, Anna Van Wersch, FF Sniehotta, and MI Trenell. Successful behavioural strategies to increase physical activity and improve glucose control in adults with type 2 diabetes. *Diabetic Medicine*, 32(8):1058–1062, 2015.

-
- [18] Franz Baader, Diego Calvanese, Deborah McGuinness, Peter Patel-Schneider, and Daniele Nardi. *The description logic handbook: Theory, implementation and applications*. Cambridge university press, 2003.
 - [19] Albert Bandura. Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, 84(2):191, 1977.
 - [20] Albert Bandura. Social cognitive theory: An agentic perspective. *Annual review of psychology*, 52(1):1–26, 2001.
 - [21] Thomas Baranowski, Cheryl L Perry, and Guy S Parcel. How individuals, environments, and health behavior interact. *Health behavior and health education: Theory, research, and practice*, 3:165–184, 2002.
 - [22] Jennifer E Barrett, Ronald C Plotnikoff, Kerry S Courneya, and Kim D Raine. Physical activity and type 2 diabetes. *The Diabetes Educator*, 33(1):128–143, 2007.
 - [23] Sotiris Batsakis and Euripides GM Petrakis. Representing temporal knowledge in the semantic web: The extended 4d fluents approach. In *Combinations of intelligent methods and applications*, pages 55–69. Springer, 2011.
 - [24] Sean Bechhofer, Frank Van Harmelen, Jim Hendler, Ian Horrocks, Deborah L McGuinness, Peter F Patel-Schneider, Lynn Andrea Stein, et al. Owl web ontology language reference. *W3C recommendation*, 10(02), 2004.
 - [25] Kathleen Benjamin and Tam Truong Donnelly. Barriers and facilitators influencing the physical activity of arabic adults: A literature review. *Avicenna*, page 8, 2013.
 - [26] Shane Bergsma and Grzegorz Kondrak. Alignment-based discriminative string similarity. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pages 656–663, 2007.
 - [27] Shlomo Berkovsky, Tsvi Kuflik, and Francesco Ricci. Mediation of user models for enhanced personalization in recommender systems. *User Modeling and User-Adapted Interaction*, 18(3):245–286, 2008.
 - [28] Amaia Bernaras. Building and reusing ontologies for electrical network applications. *Proc. of ECAI 96, 1996*, pages 298–302, 1996.

- [29] Hilary J Bethancourt, Dori E Rosenberg, Tara Beatty, and David E Arterburn. Barriers to and facilitators of physical activity program use among older adults. *Clinical medicine & research*, pages cmr–2013, 2014.
- [30] Camila Bezerra, Fred Freitas, and Filipe Santana. Evaluating ontologies with competency questions. In *2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, volume 3, pages 284–285. IEEE, 2013.
- [31] D Bhugra, A Ventriglio, MR Kuzman, G Ikkos, MH-M Hermans, P Falkai, Andrea Fiorillo, M Musalek, C Hoschl, J Dales, et al. Epa guidance on the role and responsibilities of psychiatrists. *European Psychiatry*, 30(3):417–422, 2015.
- [32] Timothy W Bickmore, Daniel Schulman, and Candace L Sidner. A reusable framework for health counseling dialogue systems based on a behavioral medicine ontology. *Journal of biomedical informatics*, 44(2):183–197, 2011.
- [33] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.
- [34] Alison O Booth, Carole Lowis, Moira Dean, Steven J Hunter, and Michelle C McKinley. Diet and physical activity in the self-management of type 2 diabetes: barriers and facilitators identified by patients and health professionals. *Primary health care research & development*, 14(3):293–306, 2013.
- [35] Willem Nico Borst. *Construction of engineering ontologies for knowledge sharing and reuse*. PhD thesis, University of Twente, Netherlands, 1997.
- [36] George Bozionelos and Paul Bennett. The theory of planned behaviour as predictor of exercise: The moderating influence of beliefs and personality variables. *Journal of health psychology*, 4(4):517–529, 1999.
- [37] Janez Brank, Marko Grobelnik, and Dunja Mladenic. A survey of ontology evaluation techniques. In *Proceedings of the conference on data mining and data warehouses (SiKDD 2005)*, pages 166–170. Citeseer Ljubljana, Slovenia, 2005.
- [38] Christopher Brewster, Harith Alani, Srinandan Dasmahapatra, and Yorick Wilks. Data driven ontology evaluation. In *International Conference on Language Resources and Evaluation, Portugal, 2004-05-24 - 2004-05-30*, 2004.

-
- [39] Dan Brickley, Ramanathan V Guha, and Andrew Layman. Resource description framework (rdf) schema specification. Technical report, W3C, 1999. W3C Proposed Recommendation, 1999.
 - [40] Jeen Broekstra, Michel Klein, Stefan Decker, Dieter Fensel, Frank Van Harmelen, and Ian Horrocks. Enabling knowledge representation on the web by extending rdf schema. In *Proceedings of the 10th international conference on World Wide Web*, pages 467–478, 2001.
 - [41] Graciela Brusa, Ma Laura Caliusco, and Omar Chiotti. A process for building a domain ontology: an experience in developing a government budgetary ontology. In *Proceedings of the second Australasian workshop on Advances in ontologies-Volume 72*, pages 7–15. Australian Computer Society, Inc., 2006.
 - [42] Michael Buckland and Fredric Gey. The relationship between recall and precision. *Journal of the American society for information science*, 45(1):12–19, 1994.
 - [43] James Cane, Denise O’Connor, and Susan Michie. Validation of the theoretical domains framework for use in behaviour change and implementation research. *Implementation science*, 7(1):37, 2012.
 - [44] Jaime Cantais, David Dominguez, Valeria Gigante, Loredana Laera, and Valentina Tamma. An example of food ontology for diabetes control. In *Proceedings of the International Semantic Web Conference 2005 workshop on Ontology Patterns for the Semantic Web*, pages 1–9, 2005.
 - [45] Carl J Caspersen, Kenneth E Powell, and Gregory M Christenson. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public health reports*, 100(2):126, 1985.
 - [46] Olivier Luc Charansonney, Luc Vanhees, and Alain Cohen-Solal. Physical activity: from epidemiological evidence to individualized patient management. *International journal of cardiology*, 170(3):350–357, 2014.
 - [47] Miao Chen. Exploring the use of ontological relations in information retrieval. In *IDEALS, Illinois Digital Environment for Access to Learning and Scholarship, iConference 2009*, 2009.

-
- [48] Ami T Choksi and Devesh C Jinwala. A novel way to relate ontology classes. *The Scientific World Journal*, 2015, 2015.
 - [49] Nicholas Christakis. *Nicholas Christakis: The hidden influence of social networks*. TED, 2010.
 - [50] Paolo Ciccarese, Marco Ocana, Leyla Jael Garcia Castro, Sudeshna Das, and Tim Clark. An open annotation ontology for science on web 3.0. *Journal of biomedical semantics*, 2(S2):S4, 2011.
 - [51] Philipp Cimiano and Johanna Völker. text2onto. In *International conference on application of natural language to information systems*, pages 227–238. Springer, 2005.
 - [52] Sheri R Colberg, Ronald J Sigal, Jane E Yardley, Michael C Riddell, David W Dunstan, Paddy C Dempsey, Edward S Horton, Kristin Castorino, and Deborah F Tate. Physical activity/exercise and diabetes: a position statement of the american diabetes association. *Diabetes care*, 39(11):2065–2079, 2016.
 - [53] Tim Coleman. Do financial incentives for delivering health promotion counselling work? analysis of smoking cessation activities stimulated by the quality and outcomes framework. *BMC public health*, 10(1):167, 2010.
 - [54] Gene Ontology Consortium. The gene ontology project in 2008. *Nucleic acids research*, 36(suppl_1):D440–D444, 2007.
 - [55] Kevin A Cradock, Gearóid ÓLaighin, Francis M Finucane, Heather L Gainforth, Leo R Quinlan, and Kathleen A Martin Ginis. Behaviour change techniques targeting both diet and physical activity in type 2 diabetes: A systematic review and meta-analysis. *International Journal of Behavioral Nutrition and Physical Activity*, 14(1):18, 2017.
 - [56] Matteo Cristani and Roberta Cuel. A survey on ontology creation methodologies. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 1(2):49–69, 2005.
 - [57] CC Curioni and PM Lourenco. Long-term weight loss after diet and exercise: a systematic review. *International journal of obesity*, 29(10):1168, 2005.
 - [58] D3C. Rdf schema. (accessed May 2020).

-
- [59] Claudia d’Amato, Nicola Fanizzi, and Floriana Esposito. A dissimilarity measure for the alc description logic. In *SWAP*, 2005.
- [60] Claudia d’Amato, Nicola Fanizzi, and Floriana Esposito. A semantic similarity measure for expressive description logics. *arXiv preprint arXiv:0911.5043*, 2009.
- [61] Andrew Darnton. An overview of behaviour change models and their uses. *Government Social Research*, 2008.
- [62] Karina W Davidson, Michael Goldstein, Robert M Kaplan, Peter G Kaufmann, Genell L Knatterud, C Tracy Orleans, Bonnie Spring, Kimberlee J Trudeau, and Evelyn P Whitlock. Evidence-based behavioral medicine: what is it and how do we achieve it? *Annals of Behavioral Medicine*, 26(3):161–171, 2003.
- [63] John Davies, Dieter Fensel, and Frank Van Harmelen. *Towards the semantic web: ontology-driven knowledge management*. John Wiley & Sons, 2003.
- [64] Philippa Davies, Anne E Walker, and Jeremy M Grimshaw. A systematic review of the use of theory in the design of guideline dissemination and implementation strategies and interpretation of the results of rigorous evaluations. *Implementation Science*, 5(1):14, 2010.
- [65] Rachel Davis, Rona Campbell, Zoe Hildon, Lorna Hobbs, and Susan Michie. Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. *Health psychology review*, 9(3):323–344, 2015.
- [66] Stefan Decker, Sergey Melnik, Frank Van Harmelen, Dieter Fensel, Michel Klein, Jeen Broekstra, Michael Erdmann, and Ian Horrocks. The semantic web: The roles of xml and rdf. *IEEE Internet computing*, 4(5):63–73, 2000.
- [67] Auriol Degbelo. A snapshot of ontology evaluation criteria and strategies. In *Proceedings of the 13th International Conference on Semantic Systems*, pages 1–8, 2017.
- [68] Ryan T Demmer, Sara Gelb, Shakira F Suglia, Katherine M Keyes, Allison E Aiello, Paolo C Colombo, Sandro Galea, Monica Uddin, Karestan C Koenen, and Laura D Kubzansky. Sex differences in the association between depression, anxiety, and type 2 diabetes mellitus. *Psychosomatic medicine*, 77(4):467, 2015.

- [69] Carlo C DiClemente, James O Prochaska, Scott K Fairhurst, Wayne F Velicer, Mary M Velasquez, and Joseph S Rossi. The process of smoking cessation: an analysis of precontemplation, contemplation, and preparation stages of change. *Journal of consulting and clinical psychology*, 59(2):295, 1991.
- [70] Loretta DiPietro. Physical activity in aging: changes in patterns and their relationship to health and function. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 56(suppl_2):13–22, 2001.
- [71] Stephan U Dombrowski, Alison Avenell, and Falko F Sniehot. Behavioural interventions for obese adults with additional risk factors for morbidity: systematic review of effects on behaviour, weight and disease risk factors. *Obesity Facts*, 3(6):377–396, 2010.
- [72] Mauro Dragoni, Tania Bailoni, Rosa Maimone, and Claudio Eccher. Helis: An ontology for supporting healthy lifestyles. In *International Semantic Web Conference*, pages 53–69. Springer, 2018.
- [73] M-C Dubé, P Valois, D Prud’Homme, SJ Weisnagel, and C Lavoie. Physical activity barriers in diabetes: development and validation of a new scale. *Diabetes research and clinical practice*, 72(1):20–27, 2006.
- [74] Martine Duclos, Sylvie Dejager, Nicolas Postel-Vinay, Sylvie di Nicola, Stéphane Quéré, and Béatrice Fiquet. Physical activity in patients with type 2 diabetes and hypertension—insights into motivations and barriers from the mobile study. *Vascular health and risk management*, 11:361, 2015.
- [75] Gareth R Dutton, Jolene Johnson, Dori Whitehead, Jamie S Bodenlos, and Phillip J Brantley. Barriers to physical activity among predominantly low-income african-american patients with type 2 diabetes. *Diabetes care*, 28(5):1209–1210, 2005.
- [76] Gunther Eysenbach, Consort-EHEALTH Group, et al. Consort-ehealth: improving and standardizing evaluation reports of web-based and mobile health interventions. *Journal of medical Internet research*, 13(4):e126, 2011.
- [77] Mariano Fernández-López, Asunción Gómez-Pérez, and Natalia Juristo. Methontology: from ontological art towards ontological engineering. In *AAAI 1997*. American Association for Artificial Intelligence, 1997.

- [78] Martin Fishbein, Harry C Triandis, Frederick H Kanfer, Marshall Becker, Susan E Middlestadt, Anita Eichler, et al. Factors influencing behavior and behavior change. *Handbook of health psychology*, 3:17, 2001.
- [79] Martin Fitzgerald and Tracy McClelland. What makes a mobile app successful in supporting health behaviour change? *Health education journal*, 76(3):373–381, 2017.
- [80] Caragh Flannery, Sheena McHugh, Ann Ebere Anaba, E Clifford, Mairead O’Riordan, Louise C Kenny, Fionnuala M McAuliffe, Patricia M Kearney, and M Byrne. Enablers and barriers to physical activity in overweight and obese pregnant women: an analysis informed by the theoretical domains framework and com-b model. *BMC pregnancy and childbirth*, 18(1):178, 2018.
- [81] Jennifer Floyd, Patricia M Zebrowski, and Gregory A Flamme. Stages of change and stuttering: A preliminary view. *Journal of Fluency Disorders*, 32(2):95–120, 2007.
- [82] Brian J Fogg. A behavior model for persuasive design. In *Proceedings of the 4th international Conference on Persuasive Technology*, page 40. ACM, 2009.
- [83] Aldo Gangemi. Ontology design patterns for semantic web content. In *International semantic web conference*, pages 262–276. Springer, 2005.
- [84] Aldo Gangemi, Carola Catenacci, Massimiliano Ciaramita, and Jos Lehmann. Modelling ontology evaluation and validation. In *European Semantic Web Conference*, pages 140–154. Springer, 2006.
- [85] Yassine Gargouri, Bernard Lefebvre, and Jean-guy Meunier. Ontology maintenance using textual analysis. In *Proc. 7TH World Multi Conference on Systemics, Cybernetics and Informatics, USA. List of Figures Figure*, volume 1, 2003.
- [86] Anca Gaston and Anita Cramp. Exercise during pregnancy: a review of patterns and determinants. *Journal of Science and Medicine in Sport*, 14(4):299–305, 2011.
- [87] Robert Gifford and Andreas Nilsson. Personal and social factors that influence pro-environmental concern and behaviour: A review. *International Journal of Psychology*, 49(3):141–157, 2014.
- [88] Kaushal Giri. Role of ontology in semantic web. *DESIDOC Journal of Library & Information Technology*, 31(2), 2011.

- [89] Georgios V Gkoutos, Paul N Schofield, and Robert Hoehndorf. The neurobehavior ontology: an ontology for annotation and integration of behavior and behavioral phenotypes. In *International review of neurobiology*, volume 103, pages 69–87. Elsevier, 2012.
- [90] Paul Glasziou, Emma Meats, Carl Heneghan, and Sasha Shepperd. What is missing from descriptions of treatment in trials and reviews? *Bmj*, 336(7659):1472–1474, 2008.
- [91] Asunción Gómez-Pérez. Some ideas and examples to evaluate ontologies. In *Proceedings the 11th Conference on Artificial Intelligence for Applications*, pages 299–305. IEEE, 1995.
- [92] Asunción Gómez-Pérez. Ontological engineering: A state of the art. *Expert Update: Knowledge Based Systems and Applied Artificial Intelligence*, 2(3):33–43, 1999.
- [93] Asunción Gómez-Pérez and Richard Benjamins. Overview of knowledge sharing and reuse components: Ontologies and problem-solving methods. In *Proceedings of the IJCAI-99 workshop on Ontologies and Problem-Solving methods (KRR5), Stockholm, Sweden*, pages 1–15, 1999.
- [94] Asunción Gómez-Pérez, Mariano Fernández-López, and Oscar Corcho. *Ontological Engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*. Springer Science & Business Media, 2006.
- [95] Thomas F Gordon and Douglas Walton. Legal reasoning with argumentation schemes. In *Proceedings of the 12th international conference on artificial intelligence and law*, pages 137–146. ACM, 2009.
- [96] Cyril Goutte and Eric Gaussier. A probabilistic interpretation of precision, recall and f-score, with implication for evaluation. In *European Conference on Information Retrieval*, pages 345–359. Springer, 2005.
- [97] Thomas R Gruber. A translation approach to portable ontology specifications. *Knowledge acquisition*, 5(2):199–220, 1993.
- [98] Thomas R Gruber. Toward principles for the design of ontologies used for knowledge sharing? *International journal of human-computer studies*, 43(5-6):907–928, 1995.

-
- [99] Michael Grüninger and Mark S Fox. Methodology for the design and evaluation of ontologies. In *Proceedings of the IJCAI-95 workshop on Basic Ontological Issues in Knowledge Sharing*, 1995.
 - [100] Nicola Guarino. Understanding, building and using ontologies. *International Journal of Human-Computer Studies*, 46(2-3):293–310, 1997.
 - [101] Nicola Guarino. *Formal ontology in information systems: Proceedings of the first international conference (FOIS'98), June 6-8, Trento, Italy*, volume 46. IOS press, 1998.
 - [102] Brian HW Guo and Yang Miang Goh. Ontology for design of active fall protection systems. *Automation in Construction*, 82:138–153, 2017.
 - [103] JA Hawley and MJ Gibala. What's new since hippocrates? preventing type 2 diabetes by physical exercise and diet. *Diabetologia*, 55(3):535–539, 2012.
 - [104] Dominik Heckmann and Antonio Krueger. A user modeling markup language (userml) for ubiquitous computing. In *International Conference on User Modeling*, pages 393–397. Springer, 2003.
 - [105] Dominik Heckmann, Tim Schwartz, Boris Brandherm, Michael Schmitz, and Margeritta von Wilamowitz-Moellendorff. Gumo—the general user model ontology. In *International Conference on User Modeling*, pages 428–432. Springer, 2005.
 - [106] Dominik Heckmann, Eric Schwarzkopf, Junichiro Mori, Dietmar Dengler, and Alexander Kröner. The user model and context ontology gumo revisited for future web 2.0 extensions. *Contexts and Ontologies: Representation and Reasoning*, pages 37–46, 2007.
 - [107] Dominikus Heckmann. *Ubiquitous user modeling*, volume 297. IOS Press, 2006.
 - [108] Valerie Heiss and Rick Petosa. Correlates of physical activity among adults with type 2 diabetes: a systematic literature review. *American Journal of Health Education*, 45(5):278–287, 2014.
 - [109] N Hex, C Bartlett, D Wright, M Taylor, and D Varley. Estimating the current and future costs of type 1 and type 2 diabetes in the uk, including direct health costs and indirect societal and productivity costs. *Diabetic Medicine*, 29(7):855–862, 2012.

-
- [110] Robin Hill. What sample size is “enough” in internet survey research. *Interpersonal Computing and Technology: An electronic journal for the 21st century*, 6(3-4):1–12, 1998.
 - [111] Pascal Hitzler, Markus Krotzsch, and Sebastian Rudolph. *Foundations of semantic web technologies*. Chapman and Hall/CRC, 2009.
 - [112] Nicola Hobbs, Alan Godfrey, Jose Lara, Linda Errington, Thomas D Meyer, Lynn Rochester, Martin White, John C Mathers, and Falko F Sniehotta. Are behavioral interventions effective in increasing physical activity at 12 to 36 months in adults aged 55 to 70 years? a systematic review and meta-analysis. *BMC medicine*, 11(1):75, 2013.
 - [113] Matthew Horridge, Simon Jupp, Georgina Moulton, Alan Rector, Robert Stevens, and Chris Wroe. A practical guide to building owl ontologies using protégé 4 and co-ode tools edition1. 2. *The university of Manchester*, 107, 2009.
 - [114] Ian Horrocks. Ontologies and the semantic web. *Communications of the ACM*, 51(12):58–67, 2008.
 - [115] Rizwan Iqbal, Masrah Azrifah Azmi Murad, Aida Mustapha, and Nurfadhlin Mohd Sharef. An analysis of ontology engineering methodologies: A literature review. *Research journal of applied sciences, engineering and technology*, 6(16):2993–3000, 2013.
 - [116] Stephen Isaac and William B Michael. *Handbook in research and evaluation: A collection of principles, methods, and strategies useful in the planning, design, and evaluation of studies in education and the behavioral sciences*. Edits publishers, 1995.
 - [117] Noah Ivers, Gro Jamtvedt, Signe Flottorp, Jane M Young, Jan Odgaard-Jensen, Simon D French, Mary Ann O’Brien, Marit Johansen, Jeremy Grimshaw, and Andrew D Oxman. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane database of systematic reviews*, 6, 2012.
 - [118] Isabel Shizu Miyamae Iwanaga, The-Minh Nguyen, Takahiro Kawamura, Hiroyuki Nakagawa, Yasuyuki Tahara, and Akihiko Ohsuga. Building an earthquake evacuation ontology from twitter. In *2011 IEEE International Conference on Granular Computing*, pages 306–311. IEEE, 2011.

- [119] Nancy K Janz and Marshall H Becker. The health belief model: A decade later. *Health education quarterly*, 11(1):1–47, 1984.
- [120] Dean Jones, Trevor Bench-Capon, and Pepijn Visser. Methodologies for ontology development. In *IFIP world computer congress*, pages 62–75, 1998.
- [121] Patricia J Jordan, Claudio R Nigg, Gregory J Norman, Joseph S Rossi, and Sonya V Benisovich. Does the transtheoretical model need an attitude adjustment?: Integrating attitude with decisional balance as predictors of stage of change for exercise. *Psychology of Sport and Exercise*, 3(1):65–83, 2002.
- [122] Aditya Kalyanpur, Bijan Parsia, Evren Sirin, Bernardo Cuenca Grau, and James Hendler. Swoop: A web ontology editing browser. *Web Semantics: Science, Services and Agents on the World Wide Web*, 4(2):144–153, 2006.
- [123] Bhaskar Kapoor and Savita Sharma. A comparative study ontology building tools for semantic web applications. *International journal of Web & Semantic Technology (IJWesT)*, 1(3):1–13, 2010.
- [124] Martin Kavalec and Vojtech Svaték. A study on automated relation labelling in ontology learning. In *Ontology Learning from Text: Methods, evaluation and applications*, pages 44–58. ed: IOS Press, 2005.
- [125] Judy Kay. ‘um: a user modelling toolkit. In *Second International user modeling workshop*, pages 11–51, 1990.
- [126] Steven H Kelder, Deanna Hoelscher, and Cheryl L Perry. How individuals, environments, and health behaviors interact. *Health behavior: Theory, research, and practice*, 159, 2015.
- [127] Michael P Kelly and Mary Barker. Why is changing health-related behaviour so difficult? *Public health*, 136:109–116, 2016.
- [128] Catriona M Kennedy, John Powell, Thomas H Payne, John Ainsworth, Alan Boyd, and Iain Buchan. Active assistance technology for health-related behavior change: an interdisciplinary review. *Journal of medical Internet research*, 14(3):e80, 2012.

-
- [129] M Rahamatullah Khondoker and Paul Mueller. Comparing ontology development tools based on an online survey. In *World Congress on Engineering 2010 (WCE 2010)*, London, UK, 2010.
- [130] Christopher SG Khoo and Jin-Cheon Na. Semantic relations in information science. *Annual review of information science and technology*, 40(1):157–228, 2006.
- [131] Warren A Kibbe, Cesar Arze, Victor Felix, Elvira Mitraka, Evan Bolton, Gang Fu, Christopher J Mungall, Janos X Binder, James Malone, Drashti Vasant, et al. Disease ontology 2015 update: an expanded and updated database of human diseases for linking biomedical knowledge through disease data. *Nucleic acids research*, 43(D1):D1071–D1078, 2014.
- [132] Panayotis Kikiras, Vassileios Tsetsos, and Stathes Hadjiefthymiades. Ontology-based user modeling for pedestrian navigation systems. In *ECAI 2006 Workshop on Ubiquitous User Modeling (UbiqUM)*, Riva del Garda, Italy, pages 1–6, 2006.
- [133] Cindy Klug, Deborah J Toobert, and Michaela Fogerty. Healthy changesTM for living with diabetes. *The Diabetes Educator*, 34(6):1053–1061, 2008.
- [134] Holger Knublauch, Matthew Horridge, Mark A Musen, Alan L Rector, Robert Stevens, Nick Drummond, Phillip W Lord, Natalya Fridman Noy, Julian Seidenberg, and Hai Wang. The protege owl experience. In *OWLED*, 2005.
- [135] Sebastian Köhler, Nicole A Vasilevsky, Mark Engelstad, Erin Foster, Julie McMurry, Ségolène Aymé, Gareth Baynam, Susan M Bello, Cornelius F Boerkoel, Kym M Boycott, et al. The human phenotype ontology in 2017. *Nucleic acids research*, 45(D1):D865–D876, 2016.
- [136] Eveliina Korkiakangas, Anja M Taanila, and Sirkka Keinänen-Kiukaanniemi. Motivation to physical activity among adults with high risk of type 2 diabetes who participated in the oulu substudy of the finnish diabetes prevention study. *Health & social care in the community*, 19(1):15–22, 2011.
- [137] Eveliina E Korkiakangas, Maija A Alahuhta, Päivi M Husman, Sirkka Keinänen-Kiukaanniemi, Anja M Taanila, and Jaana H Laitinen. Motivators and barriers to exercise among adults with a high risk of type 2 diabetes—a qualitative study. *Scandinavian journal of caring sciences*, 25(1):62–69, 2011.

-
- [138] Egor V Kostylev, Juan L Reutter, and Martín Ugarte. Construct queries in sparql. In *18th International Conference on Database Theory (ICDT 2015)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2015.
- [139] Dominika Kwasnicka, Stephan U Dombrowski, Martin White, and Falko Sniehotta. Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. *Health psychology review*, 10(3):277–296, 2016.
- [140] Kai R Larsen, Susan Michie, Eric B Hekler, Bryan Gibson, Donna Spruijt-Metz, David Ahern, Heather Cole-Lewis, Rebecca J Bartlett Ellis, Bradford Hesse, Richard P Moser, et al. Behavior change interventions: the potential of ontologies for advancing science and practice. *Journal of behavioral medicine*, 40(1):6–22, 2017.
- [141] Ora Lassila and Deborah McGuinness. The role of frame-based representation on the semantic web. *Linköping Electronic Articles in Computer and Information Science*, 6(5):2001, 2001.
- [142] Julia Lawton, N Ahmad, L Hanna, M Douglas, and N Hallowell. ?i can’t do any serious exercise?: barriers to physical activity amongst people of pakistani and indian origin with type 2 diabetes. *Health education research*, 21(1):43–54, 2006.
- [143] LP Lidegaard, Nete Schwennesen, Ingrid Willaing, and Kristine Færch. Barriers to and motivators for physical activity among people with type 2 diabetes: patients’ perspectives. *Diabetic Medicine*, 33(12):1677–1685, 2016.
- [144] Angela D Liese, Xiaoguang Ma, David M Maahs, and Jennifer L Trilk. Physical activity, sedentary behaviors, physical fitness, and their relation to health outcomes in youth with type 1 and type 2 diabetes: A review of the epidemiologic literature. *Journal of Sport and Health Science*, 2(1):21–38, 2013.
- [145] Edward Loper and Steven Bird. Nltk: the natural language toolkit. *arXiv preprint cs/0205028*, 2002.
- [146] Mariano Fernández López, Asunción Gómez-Pérez, Juan Pazos Sierra, and Alejandro Pazos Sierra. Building a chemical ontology using methontology and the ontology design environment. *IEEE Intelligent Systems and their applications*, 14(1):37–46, 1999.

- [147] Fabiana Lorencatto, Robert West, Carla Bruguera, Leonie S Brose, and Susan Michie. Assessing the quality of goal setting in behavioural support for smoking cessation and its association with outcomes. *Annals of Behavioral Medicine*, 50(2):310–318, 2015.
- [148] Adolfo Lozano-Tello and Asunción Gómez-Pérez. Ontometric: A method to choose the appropriate ontology. *Journal of Database Management (JDM)*, 15(2):1–18, 2004.
- [149] Gay L.R and Diehl P.L. Research methods for business and management. *New York: Macmillan.*, 1992.
- [150] David R Lubans, Ronald C Plotnikoff, Mary Jung, Neil Eves, and Ron Sigal. Testing mediator variables in a resistance training intervention for obese adults with type 2 diabetes. *Psychology & Health*, 27(12):1388–1404, 2012.
- [151] Alexander Maedche and Steffen Staab. Measuring similarity between ontologies. In *International Conference on Knowledge Engineering and Knowledge Management*, pages 251–263. Springer, 2002.
- [152] Alexander Maedche and Raphael Volz. The ontology extraction & maintenance framework text-to-onto. In *Proc. Workshop on Integrating Data Mining and Knowledge Management, USA*, pages 1–12, 2001.
- [153] Golemati Maria, Katifori Akrivi, Vassilakis Costas, L George, and Halatsis Constantin. Creating an ontology for the user profile: Method and applications. In *Proceedings AI* AI Workshop RCIS*, 2007.
- [154] Emma S Martin, Fiona Dobson, Michelle Hall, Charlotte Marshall, and Thorlene Egerton. The effects of behavioural counselling on the determinants of health behaviour change in adults with chronic musculoskeletal conditions making lifestyle changes: A systematic review and meta-analysis. *Musculoskeletal care*, 17(3):170–197, 2019.
- [155] Nicola McCleary, Eilidh M Duncan, Fiona Stewart, and Jill J Francis. Active ingredients are reported more often for pharmacologic than non-pharmacologic interventions: an illustrative review of reporting practices in titles and abstracts. *Trials*, 14(1):146, 2013.
- [156] Deborah L McGuinness. Ontologies come of age. *Spinning the semantic web: bringing the World Wide Web to its full potential*, pages 171–194, 2002.

-
- [157] Amanda M McGuire, Debra J Anderson, and Paul Fulbrook. Perceived barriers to healthy lifestyle activities in midlife and older australian women with type 2 diabetes. *Collegian*, 21(4):301–310, 2014.
- [158] D. McKenzie-Mohr. Promoting sustainable behavior: An introduction to community-based social marketing. *Journal of Social Issues*, 56:543–554, 01 2000.
- [159] Arjuna Medagama and Manoj Galgomuwa. Comorbidities and ethnocultural factors limit the physical activity of rural sri lankan patients with diabetes mellitus. *Journal of diabetes research*, 2018, 2018.
- [160] Susan Michie and Charles Abraham. Advancing the science of behaviour change: A plea for scientific reporting. *Addiction*, 103(9):1409–1410, 2008.
- [161] Susan Michie, Charles Abraham, Martin P Eccles, Jill J Francis, Wendy Hardeman, and Marie Johnston. Strengthening evaluation and implementation by specifying components of behaviour change interventions: a study protocol. *Implementation Science*, 6(1):10, 2011.
- [162] Susan Michie, Charles Abraham, Craig Whittington, John McAteer, and Sunjai Gupta. Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychology*, 28(6):690, 2009.
- [163] Susan Michie, Stefanie Ashford, Falko F Sniehotta, Stephan U Dombrowski, Alex Bishop, and David P French. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the calo-re taxonomy. *Psychology & health*, 26(11):1479–1498, 2011.
- [164] Susan Michie, Lou Atkins, and Robert West. The behaviour change wheel. *A guide to designing interventions. 1st ed. Great Britain: Silverback Publishing*, pages 1003–1010, 2014.
- [165] Susan Michie, Rachel N Carey, Marie Johnston, Alexander J Rothman, Marijn De Bruin, Michael P Kelly, and Lauren E Connell. From theory-inspired to theory-based interventions: A protocol for developing and testing a methodology for linking behaviour change techniques to theoretical mechanisms of action. *Annals of behavioral medicine*, 52(6):501–512, 2018.

-
- [166] Susan Michie, Marie Johnston, Jill Francis, Wendy Hardeman, and Martin Eccles. From theory to intervention: mapping theoretically derived behavioural determinants to behaviour change techniques. *Applied psychology*, 57(4):660–680, 2008.
- [167] Susan Michie and Andrew Prestwich. Are interventions theory-based? development of a theory coding scheme. *Health psychology*, 29(1):1, 2010.
- [168] Susan Michie, Michelle Richardson, Marie Johnston, Charles Abraham, Jill Francis, Wendy Hardeman, Martin P Eccles, James Cane, and Caroline E Wood. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Annals of behavioral medicine*, 46(1):81–95, 2013.
- [169] Susan Michie, James Thomas, Marie Johnston, Pol Mac Aonghusa, John Shawe-Taylor, Michael P Kelly, Léa A Deleris, Ailbhe N Finnerty, Marta M Marques, Emma Norris, et al. The human behaviour-change project: harnessing the power of artificial intelligence and machine learning for evidence synthesis and interpretation. *Implementation Science*, 12(1):121, 2017.
- [170] Susan Michie, Maartje M Van Stralen, and Robert West. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation science*, 6(1):42, 2011.
- [171] Susan Michie, Caroline E Wood, Marie Johnston, Charles Abraham, Jill Francis, and Wendy Hardeman. Behaviour change techniques: the development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (a suite of five studies involving consensus methods, randomised controlled trials and analysis of qualitative data). *Health Technology Assessment*, 19(99), 2015.
- [172] Nelda Mier, Alvaro A Medina, and Marcia G Ory. Peer reviewed: Mexican americans with type 2 diabetes: Perspectives on definitions, motivators, and programs of physical activity. *Preventing chronic disease*, 4(2), 2007.
- [173] George A Miller. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995.
- [174] Stephania T Miller, Khensani N Marolen, and Bettina M Beech. Perceptions of

- physical activity and motivational interviewing among rural african-american women with type 2 diabetes. *Women's health issues*, 20(1):43–49, 2010.
- [175] Verywell Mind. The role and training of a clinical psychologist. *available at: <https://www.verywellmind.com/clinical-psychologist-2510503>*, 2020. (accessed August 2020).
- [176] Riichiro Mizoguchi. Tutorial on ontological engineering part 2: Ontology development, tools and languages. *New Generation Computing*, 22(1):61–96, 2004.
- [177] Riichiro Mizoguchi and Mitsuru Ikeda. Towards ontology engineering. *Journal-Japanese Society for Artificial Intelligence*, 13:9–10, 1998.
- [178] Riichiro Mizoguchi, Johan Vanwelkenhuysen, and Mitsuru Ikeda. Task ontology for reuse of problem solving knowledge. *Towards Very Large Knowledge Bases: Knowledge Building & Knowledge Sharing*, 46(59):45, 1995.
- [179] David Moher, Kenneth F Schulz, Douglas G Altman, Consort Group, et al. The consort statement: revised recommendations for improving the quality of reports of parallel-group randomised trials, 2001.
- [180] David C Mohr, Stephen M Schueller, Enid Montague, Michelle Nicole Burns, and Parisa Rashidi. The behavioral intervention technology model: an integrated conceptual and technological framework for ehealth and mhealth interventions. *Journal of medical Internet research*, 16(6), 2014.
- [181] Graham F Moore, Suzanne Audrey, Mary Barker, Lyndal Bond, Chris Bonell, Wendy Hardeman, Laurence Moore, Alicia O’Cathain, Tannaze Tinati, Daniel Wight, et al. Process evaluation of complex interventions: Medical research council guidance. *bmj*, 350:h1258, 2015.
- [182] Elaine H Morrato, James O Hill, Holly R Wyatt, Vahram Ghushchyan, and Patrick W Sullivan. Physical activity in us adults with diabetes and at risk for developing diabetes, 2003. *Diabetes care*, 30(2):203–209, 2007.
- [183] Patrick J Morris. Physical activity recommendations for children and adolescents with chronic disease. *Current Sports Medicine Reports*, 7(6):353–358, 2008.

-
- [184] Kamran Munir and M Sheraz Anjum. The use of ontologies for effective knowledge modelling and information retrieval. *Applied Computing and Informatics*, 14(2):116–126, 2018.
 - [185] David M Nathan. Initial management of glycemia in type 2 diabetes mellitus. *New England Journal of Medicine*, 347(17):1342–1349, 2002.
 - [186] Robert Neches, Richard E Fikes, Tim Finin, Thomas Gruber, Ramesh Patil, Ted Senator, and William R Swartout. Enabling technology for knowledge sharing. *AI magazine*, 12(3):36, 1991.
 - [187] Lillian M Nejad, Eleanor H Wertheim, and Ken Greenwood. Comparison of the health belief model and the theory of planned behavior in the prediction of dieting and fasting behavior. *Sensoria: A Journal of Mind, Brain & Culture*, 1(1):63–74, 2005.
 - [188] Seth M Noar and Rick S Zimmerman. Health behavior theory and cumulative knowledge regarding health behaviors: are we moving in the right direction? *Health education research*, 20(3):275–290, 2005.
 - [189] Emma Norris, Ailbhe N Finnerty, Janna Hastings, Gillian Stokes, and Susan Michie. A scoping review of ontologies related to human behaviour change. *Nature Human Behaviour*, page 1, 2019.
 - [190] Natalya Noy, Tania Tudorache, Csongor Nyulas, and Mark Musen. The ontology life cycle: Integrated tools for editing, publishing, peer review, and evolution of ontologies. In *AMIA Annual Symposium Proceedings*, volume 2010, page 552. American Medical Informatics Association, 2010.
 - [191] Natalya F Noy, Abhita Chugh, William Liu, and Mark A Musen. A framework for ontology evolution in collaborative environments. In *International semantic web conference*, pages 544–558. Springer, 2006.
 - [192] Natalya F Noy, Deborah L McGuinness, et al. *Ontology development 101: A guide to creating your first ontology*, 2001.
 - [193] Natalya Fridman Noy and Carole D Hafner. The state of the art in ontology design: A survey and comparative review. *AI magazine*, 18(3):53–53, 1997.

-
- [194] Natasha Noy, Alan Rector, Pat Hayes, and Chris Welty. Defining n-ary relations on the semantic web. *W3C working group note*, 12(4), 2006.
- [195] Pa Pa Nyunt and Ni Lar Thein. Software agent oriented information integration system in semantic web. In *Information and Telecommunication Technologies, 2005. APSITT 2005 Proceedings. 6th Asia-Pacific Symposium on*, pages 266–271. IEEE, 2005.
- [196] Leo Obrst, Werner Ceusters, Inderjeet Mani, Steve Ray, and Barry Smith. The evaluation of ontologies. In *Semantic web*, pages 139–158. Springer, 2007.
- [197] United States National Library of Medicine (NLM). Medical subject headings, 2019.
- [198] James M Olson. Psychological barriers to behavior change: How to indentify the barriers that inhibit change. *Canadian Family Physician*, 38:309, 1992.
- [199] World Health Organization. *How to use the ICF. A Practical Manual for using the International Classification of Functioning, Disability and Health (ICF). Exposure draft for comment.*, October 2013. Geneva. Retrieved from: <http://www.who.int/classifications/drafticfpracticalmanual.pdf>.
- [200] World Health Organization et al. *Global report on diabetes*. World Health Organization, 2016.
- [201] Thearis Osuji, Sarah Lovegreen, Michael Elliott, and Ross C Brownson. Barriers to physical activity among women in the rural midwest. *Women & Health*, 44(1):41–55, 2006.
- [202] Una-May O’Reilly and Martin Hemberg. Integrating generative growth and evolutionary computation for form exploration. *Genetic Programming and Evolvable Machines*, 8(2):163–186, 2007.
- [203] Camille A Pearte, Tiffany L Gary, and Frederick L Brancati. Correlates of physical activity levels in a sample of urban african americans with type 2 diabetes. *Ethnicity and Disease*, 14(2):198–205, 2004.
- [204] Nola J Pender, Carolyn L Murdaugh, Mary Ann Parsons, et al. *Health promotion in nursing practice*. Prentice Hall Upper Saddle River, NJ, 2006.

- [205] Constance Percy, Valerie van Holten, and Calum S Muir. *International classification of diseases for oncology*. Geneva: World Health Organization, 1990.
- [206] Nhathai Phan, Dejing Dou, Hao Wang, David Kil, and Brigitte Piniewski. Ontology-based deep learning for human behavior prediction with explanations in health social networks. *Information sciences*, 384:298–313, 2017.
- [207] Ronald C Plotnikoff, Sonia Lippke, Linda Trinh, Kerry S Courneya, Nick Birkett, and Ronald J Sigal. Protection motivation theory and the prediction of physical activity among adults with type 1 or type 2 diabetes in a large population sample. *British Journal of Health Psychology*, 15(3):643–661, 2010.
- [208] Robert Porzel and Rainer Malaka. A task-based approach for ontology evaluation. In *ECAI Workshop on Ontology Learning and Population, Valencia, Spain*, pages 1–6. Citeseer, 2004.
- [209] Djoko Pramono, Nanang Yudi Setiawan, Riyanarto Sarno, and Mohamad Sidiq. Physical activity recommendation for diabetic patients based on ontology. In *7th International Conference on Information & Communication Technology and Systems*, pages 27–32, 2013.
- [210] Alun Preece, Mario Gomez, Geeth De Mel, Wamberto Vasconcelos, Derek Sleeman, Stuart Colley, and Tom La Porta. An ontology-based approach to sensor-mission assignment. In *1st Annual Conference of the International Technology Alliance (ACITA), Maryland, USA*, page 30. Citeseer, 2007.
- [211] James O Prochaska and Bess H Marcus. The transtheoretical model: Applications to exercise. In RK Dishman, editor, *Advances in exercise adherence*, pages 161–180. Human Kinetics Publishers, 1994.
- [212] James O Prochaska and John C Norcross. Stages of change. *Psychotherapy: theory, research, practice, training*, 38(4):443, 2001.
- [213] James O Prochaska and Wayne F Velicer. The transtheoretical model of health behavior change. *American journal of health promotion*, 12(1):38–48, 1997.
- [214] Eric Prud’hommeaux and Andy Seaborne. Sparql query language for rdf. w3c working draft. *World Wide Web Consortium, February. Latest version: <http://www.w3.org/TR/rdf-sparql-query>*, 2006.

- [215] Matthew Purver, Thomas L Griffiths, Konrad P Körding, and Joshua B Tenenbaum. Unsupervised topic modelling for multi-party spoken discourse. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*, pages 17–24. Association for Computational Linguistics, 2006.
- [216] Lyn Quine, Derek R Rutter, and Laurence Arnold. Predicting and understanding safety helmet use among schoolboy cyclists: a comparison of the theory of planned behaviour and the health belief model. *Psychology and Health*, 13(2):251–269, 1998.
- [217] Dnyanesh Rajpathak, Enrico Motta, and Rajkumar Roy. A generic task ontology for scheduling applications. In *International Conference on Artificial Intelligence (IC AI'2001)*, 2001.
- [218] Martin Raubal and Werner Kuhn. Ontology-based task simulation. *Spatial Cognition and Computation*, 4(1):15–37, 2004.
- [219] Liana Razmerita, Albert Angehrn, and Alexander Maedche. Ontology-based user modeling for knowledge management systems. In *International Conference on User Modeling*, pages 213–217. Springer, 2003.
- [220] Alan Rector, Nick Drummond, Matthew Horridge, Jeremy Rogers, Holger Knublauch, Robert Stevens, Hai Wang, and Chris Wroe. Owl pizzas: Practical experience of teaching owl-dl: Common errors & common patterns. In *International Conference on Knowledge Engineering and Knowledge Management*, pages 63–81. Springer, 2004.
- [221] Mel Reed and Bev Lloyd. *Health psychology*. Scientific e-Resources, 2018.
- [222] Caroline R Richardson, Kathleen S Mehari, Laura G McIntyre, Adrienne W Janney, Laurie A Fortlage, Ananda Sen, Victor J Strecher, and John D Piette. A randomized trial comparing structured and lifestyle goals in an internet-mediated walking program for people with type 2 diabetes. *International Journal of Behavioral Nutrition and Physical Activity*, 4(1):59, 2007.
- [223] Eric Sven Ristad and Peter N Yianilos. Learning string-edit distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(5):522–532, 1998.

-
- [224] Julia Rogushina and Anatoly Gladun. Ontology-based competency analyses in new research domains. *Journal of Computing and Information Technology*, 20(4):277–291, 2012.
- [225] Monica Roman, Ioana Sandu, and Sabin C Buraga. Owl-based modeling of rpg games. *Studia Universitatis Babes-Bolyai*, 56(3):83, 2011.
- [226] IM Rosenstock and JP Kirscht. The health belief model and personal health behaviour. *Health Education Monogr*, 2:470–473, 1974.
- [227] Bart PF Rutten, C Hammels, N Geschwind, C Menne-Lothmann, E Pishva, K Schruers, D Van Den Hove, G Kenis, J Van Os, and M Wichers. Resilience in mental health: linking psychological and neurobiological perspectives. *Acta Psychiatrica Scandinavica*, 128(1):3–20, 2013.
- [228] Marta Sabou and Miriam Fernandez. Ontology (network) evaluation. In *Ontology engineering in a networked world*, pages 193–212. Springer, 2012.
- [229] Carmen D Samuel-Hodge, Sandra W Headen, Anne H Skelly, Allyson F Ingram, Thomas C Keyserling, Ethel J Jackson, Alice S Ammerman, and Tom A Elasy. Influences on day-to-day self-management of type 2 diabetes among african-american women: spirituality, the multi-caregiver role, and other social context factors. *Diabetes care*, 23(7):928–933, 2000.
- [230] Lynn Marie Schriml, Cesar Arze, Suvarna Nadendla, Yu-Wei Wayne Chang, Mark Mazaitis, Victor Felix, Gang Feng, and Warren Alden Kibbe. Disease ontology: a backbone for disease semantic integration. *Nucleic acids research*, 40(D1):D940–D946, 2011.
- [231] Lynn Marie Schriml, Cesar Arze, Suvarna Nadendla, Yu-Wei Wayne Chang, Mark Mazaitis, Victor Felix, Gang Feng, and Warren Alden Kibbe. Disease ontology: a backbone for disease semantic integration. *Nucleic acids research*, 40(D1):D940–D946, 2012.
- [232] Karen A Schutzer and B Sue Graves. Barriers and motivations to exercise in older adults. *Preventive medicine*, 39(5):1056–1061, 2004.

-
- [233] Sascha Seifert, Michael Kelm, Manuel Moeller, Saikat Mukherjee, Alexander Cavallaro, Martin Huber, and Dorin Comaniciu. Semantic annotation of medical images. In *Medical Imaging 2010: Advanced PACS-based Imaging Informatics and Therapeutic Applications*, volume 7628, page 762808. International Society for Optics and Photonics, 2010.
- [234] THE U.S. PUPRIC HEALTH SERVICE. Commissioned corps of the u.s. pupric health service.
- [235] Jonathan E Shaw, Richard A Sicree, and Paul Z Zimmet. Global estimates of the prevalence of diabetes for 2010 and 2030. *Diabetes research and clinical practice*, 87(1):4–14, 2010.
- [236] Saurabh RamBihariLal Shrivastava, Prateek Saurabh Shrivastava, and Jegadeesh Ramasamy. Role of self-care in management of diabetes mellitus. *Journal of diabetes and Metabolic disorders*, 12(1):14, 2013.
- [237] Ronald J Sigal, Marni J Armstrong, Simon L Bacon, Normand G Boule, Kaberi Dasgupta, Glen P Kenny, and Michael C Riddell. Physical activity and diabetes. *Canadian journal of diabetes*, 42:S54–S63, 2018.
- [238] Nataliya Sklyar. Survey of existing bio-ontologies. *Institute for informatik*, 2001.
- [239] Kevin J Smith, Linda Mezydlo Subich, and Cynthia Kalodner. The transtheoretical model’s stages and processes of change and their relation to premature termination. *Journal of counseling psychology*, 42(1):34, 1995.
- [240] Misoon Song, Mira Lee, and Bomin Shim. Barriers to and facilitators of self-management adherence in korean older adults with type 2 diabetes. *International journal of older people nursing*, 5(3):211–218, 2010.
- [241] Donna Spruijt-Metz, Eric Hekler, Niilo Saranummi, Stephen Intille, Ilkka Korhonen, Wendy Nilsen, Daniel E Rivera, Bonnie Spring, Susan Michie, David A Asch, et al. Building new computational models to support health behavior change and maintenance: new opportunities in behavioral research. *Translational behavioral medicine*, 5(3):335–346, 2015.
- [242] Steffen Staab, Rudi Studer, H-P Schnurr, and York Sure. Knowledge processes and ontologies. *IEEE Intelligent systems*, 16(1):26–34, 2001.

- [243] Rudi Studer, V Richard Benjamins, Dieter Fensel, et al. Knowledge engineering: principles and methods. *Data and knowledge engineering*, 25(1):161–198, 1998.
- [244] Mari Carmen Suárez-Figueroa, Asunción Gómez-Pérez, and Boris Villazón-Terrazas. How to write and use the ontology requirements specification document. In *OTM Confederated International Conferences "On the Move to Meaningful Internet Systems"*, pages 966–982. Springer, 2009.
- [245] Boontawee Suntisrivaraporn. A similarity measure for the description logic el with unfoldable terminologies. In *2013 5th International Conference on Intelligent Networking and Collaborative Systems*, pages 408–413. IEEE, 2013.
- [246] Takeo Suzuki, Ryoko Takei, Toyoshi Inoguchi, Noriyuki Sonoda, Shuji Sasaki, Toshihiko Kaise, and Ryoichi Takayanagi. Clinical significance of barriers to blood glucose control in type 2 diabetes patients with insufficient glycemic control. *Patient preference and adherence*, 9:837, 2015.
- [247] Bill Swartout, Ramesh Patil, Kevin Knight, and Tom Russ. Toward distributed use of large-scale ontologies. In *Proc. of the Tenth Workshop on Knowledge Acquisition for Knowledge-Based Systems*, pages 138–148, 1996.
- [248] Shane N Sweet, Michelle S Fortier, Eva Guerin, Heather Tulloch, Ronald J Sigal, Glen P Kenny, and Robert D Reid. Understanding physical activity in adults with type 2 diabetes after completing an exercise intervention trial: A mediation model of self-efficacy and autonomous motivation. *Psychology, health & medicine*, 14(4):419–429, 2009.
- [249] Abraham Tamirat, Lakew Abebe, and Getachew Kirose. Prediction of physical activity among type-2 diabetes patients attending jimma university specialized hospital, southwest ethiopia: Application of health belief model. *Science*, 2(6):524–31, 2014.
- [250] Samir Tartir, I Budak Arpinar, and Amit P Sheth. Ontological evaluation and validation. In *Theory and applications of ontology: Computer applications*, pages 115–130. Springer, 2010.
- [251] Mohsen Tavakol and John Sandars. Quantitative and qualitative methods in medical education research: Amee guide no 90: Part ii. *Medical teacher*, 36(10):838–848, 2014.

-
- [252] Sabine Timpf. Ontologies of wayfinding: a traveler's perspective. *Networks and spatial economics*, 2(1):9–33, 2002.
- [253] Michael Uschold and Martin King. *Towards a methodology for building ontologies*. Citeseer, 1995.
- [254] Mike Uschold. Building ontologies: Towards a unified methodology. In *Proceedings of 16th Annual Conference of the British Computer Society Specialists Group on Expert Systems*. Citeseer, 1996.
- [255] IJM Van den Arend, RP Stolk, HMJ Krans, DE Grobbee, and AJP Schrijvers. Management of type 2 diabetes: a challenge for patient and physician. *Patient education and counseling*, 40(2):187–194, 2000.
- [256] Gertjan Van Heijst, A Th Schreiber, and Bob J Wielinga. Using explicit ontologies in kbs development. *International journal of human-computer studies*, 46(2-3):183–292, 1997.
- [257] Ben Van Ommen, Suzan Wopereis, Pepijn van Empelen, Hilde M van Keulen, Wilma Otten, Marise Kasteleyn, Johanna JW Molema, Iris M de Hoogh, Niels H Chavannes, Mattijs E Numans, et al. From diabetes care to diabetes cure—the integration of systems biology, ehealth, and behavioral change. *Frontiers in endocrinology*, 8:381, 2018.
- [258] C.J. Van Rijsbergen. *Information Retrieval*. Butterworths, 1979.
- [259] Michaël Verdonck, Frederik Gailly, Sergio de Cesare, and Geert Poels. Ontology-driven conceptual modeling: A systematic literature mapping and review. *Applied Ontology*, 10(3-4):197–227, 2015.
- [260] Pepijn RS Visser and Trevor JM Bench-Capon. A comparison of four ontologies for the design of legal knowledge systems. *Artificial Intelligence and Law*, 6(1):27–57, 1998.
- [261] Denny Vrandečić. Ontology evaluation. In *Handbook on ontologies*, pages 293–313. Springer, 2009.
- [262] W3C. Defining n-ary relations on the semantic web, 2066.

- [263] Diabetes WA. How do endocrinologists help people living with diabetes? *available at: <https://www.diabeteswa.com.au/how-do-endocrinologists-help-people-living-with-diabetes/>*, 2019. (accessed August 2020).
- [264] Dian Wang, WH Tang, and QH Wu. Ontology-based fault diagnosis for power transformers. In *IEEE PES General Meeting*, pages 1–8. IEEE, 2010.
- [265] Hao Wang, Dejing Dou, and Daniel Lowd. Ontology-based deep restricted boltzmann machine. In *Proceedings, Part I, 27th International Conference on Database and Expert Systems Applications - Volume 9827*, DEXA 2016, page 431–445, Berlin, Heidelberg, 2016. Springer-Verlag.
- [266] Xiao Hang Wang, D Qing Zhang, Tao Gu, and Hung Keng Pung. Ontology based context modeling and reasoning using owl. In *Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second IEEE Annual Conference on*, pages 18–22. Ieee, 2004.
- [267] Yiwen Wang, Natalia Stash, Lora Aroyo, Peter Gorgels, Lloyd Rutledge, and Guus Schreiber. Recommendations based on semantically enriched museum collections. *Web Semantics: Science, Services and Agents on the World Wide Web*, 6(4):283–290, 2008.
- [268] John R Weisz, Mei Yi Ng, and Sarah Kate Bearman. Odd couple? reenvisioning the relation between science and practice in the dissemination-implementation era. *Clinical Psychological Science*, 2(1):58–74, 2014.
- [269] Lonnie K Wen, Marvin D Shepherd, and Michael L Parchman. Family support, diet, and exercise among older mexican americans with type 2 diabetes. *The Diabetes Educator*, 30(6):980–993, 2004.
- [270] Robert West and Susan Michie. *A Guide to Development and Evaluation of Digital Behaviour Interventions in Healthcare*. Silverback Publishing, 2016.
- [271] Robert West, Asha Walia, Natasha Hyder, Lion Shahab, and Susan Michie. Behavior change techniques used by the english stop smoking services and their associations with short-term quit outcomes. *Nicotine & Tobacco Research*, 12(7):742–747, 2010.

-
- [272] James D Westaby. Behavioral reasoning theory: Identifying new linkages underlying intentions and behavior. *Organizational behavior and human decision processes*, 98(2):97–120, 2005.
- [273] Nancy Wiegand and Cassandra García. A task-based ontology approach to automate geospatial data retrieval. *Transactions in GIS*, 11(3):355–376, 2007.
- [274] Dawid Wisniewski. Automatic translation of competency questions into sparql-owl queries. In *Companion of the The Web Conference 2018 on The Web Conference 2018*, pages 855–859. International World Wide Web Conferences Steering Committee, 2018.
- [275] Caroline E Wood, Wendy Hardeman, Marie Johnston, Jill Francis, Charles Abraham, and Susan Michie. Reporting behaviour change interventions: do the behaviour change technique taxonomy v1, and training in its use, improve the quality of intervention descriptions? *Implementation Science*, 11(1):84, 2015.
- [276] Caroline E Wood, Michelle Richardson, Marie Johnston, Charles Abraham, Jill Francis, Wendy Hardeman, and Susan Michie. Applying the behaviour change technique (bct) taxonomy v1: a study of coder training. *Translational behavioral medicine*, 5(2):134–148, 2015.
- [277] Dell Zhang, Jun Wang, and Xiaoxue Zhao. Estimating the uncertainty of average f1 scores. In *Proceedings of the 2015 International Conference on The Theory of Information Retrieval*, pages 317–320, 2015.
- [278] Qian Zhong, Hanyu Li, Juanzi Li, Guotong Xie, Jie Tang, Lizhu Zhou, and Yue Pan. A gauss function based approach for unbalanced ontology matching. In *Proceedings of the 2009 ACM SIGMOD International Conference on Management of data*, pages 669–680. ACM, 2009.
- [279] Sonja Zillner. Towards the ontology-based classification of lymphoma patients using semantic image annotations. In *SWAT4LS*. Citeseer, 2009.

Appendix A

Systematic Review

Table A.1: The main classification of identified barriers, with supporting source quotations

Barrier	Source Quotation
Health Barrier	
Hypoglycemia or low sugar level	...potential health issues and training on potential risks such as hypoglycemia...[74].
Medical condition other than T2D (e.g. heart attack)	...medical conditions, such as asthma and heart disease...[12]. ...Clinical characteristics/comorbidities (heart disease, hypertension, arthritis) [203]. ...medical outcome variables (e.g., blood pressure, HbA1c, BMI) [75].
Environmental Barrier	
Bad weather condition	...environment influences included..., weather,...[22].
Lack of safety (e.g. gang activity)	...environmental considerations such as...safety... [143]. For some participants,...gang activity,...were environmental barriers [172]. Environment...Street safety-women felt unsafe to walk alone in the city [25].

Continued on next page

Table A.1 – *Continued from previous page*

Barrier	Source Quotation
Lack of facility or equipment	For some participants,...lack of facilities were environmental barriers [172]. ...environment influences included availability of facilities,... [22].
Facilities and equipment cost	- Common obstacles include...,environmental considerations such as cost...[143].
Transportation problem	For some participants, traffic..., and transportation were environmental barriers [172].
Personal Barrier	
Lack of knowledge	Individual...information about the benefits of exercise.. [25]. Interpersonal level factors...(e.g. lack of guidance from a professional) [29].
Lack of time	Personal barriers....lack of time: there is no time because there is a lot of housework [12].
Old age	Individual...(e.g. old age) [25].
Dislike, lack of enjoyment or lack of interest	Other important intrapersonal factors, particularly regarding the use of PA programs, included personal preferences, likes/ dislikes, intimidation, and personal history with PA. For example, some participants did not like gyms and/or group activities [29].
Financial problem	Individual...(e.g. lack of money) [25].
Laziness	Individual....Laziness.. [25].
Preference for other activities	Other important intrapersonal factors, particularly regarding the use of PA programs, included personal preferences... [29].
Physical Barrier	
Pain (e.g. injury)	Physical...(e.g. pain,...risk of injury) [29].
Fatigue and tiredness	...physical symptoms of fatigue, hip/knee pain, or back pain [203].

Continued on next page

Table A.1 – *Continued from previous page*

Barrier	Source Quotation
Psychological Barrier	
Lack of self-efficacy	Psychosocial factors included self-efficacy... [108].
Feeling depressed	Psychosocial Factors:...Affective disorders such as...depression tend to be inversely associated with physical activity participation at any age (50-52) [70].
Feeling stressed or anxious	Psychosocial Factors:....Characteristics such as, stress... [70]. psychosocial factors (...anxiety)...[240]
Social Barrier	
Lack of support or motivation	Barriers: cultural/social: Another barrier reported...lack of social support for exercise... [25].
Cultural norms	Barriers: cultural/social: Cultural norms and expectations regarding women's roles were also viewed as barriers to physical activity [25].
Embarrassment	Social and Cultural: The females said they were embarrassed [159]. Socially, embarrassment,... were common reasons for not engaging in PA [159].

Table A.2: The main classification (Clas.) of identified barriers to physical activities (PAs), with supporting source quotations

Barrier / Clas.	Prevented PAs	Source Quotation
Health Barrier		
Hypoglycemia or low sugar level	Different physical activities	- ...hypoglycemia has previously been described as a significant barrier to PA [74].
		Exercising may cause too low blood sugar [203].
		- Furthermore, some participants mentioned that having diabetes, specifically trying to balance sugar levels, could act as a hindrance [22].

Continued on next page

Barrier / Clas.	Prevented PAs	Source Quotation
Medical condition other than T2D (e.g. asthma, heart rate, breathlessness, peripheral neuropathy, degenerative joint disease, arthritis and coronary artery)	Different physical activities (e.g. weights or the machines and all that)	<p>- Many respondents' accounts were peppered with references to health complaints, such as asthma... [142].</p> <p>Furthermore, rather than seeing sweating, increased heart rate and breathlessness as 'normal' byproducts of physical activity, some respondents perceived them as illness states and thus as something they should try to avoid [142]. - I can't do any serious exercise on the weights or the machines and all that because I feel...I start feeling ill y' know if I exercise [142].</p> <p>- I can't walk. I walk a short distance and become breathless and when I become breathless, one can fall that's another problem. If you have the courage and strength then you can do everything. When I stand I become dizzy [142]. - African- American focus groups reported numerous barriers, including...peripheral neuropathy, and degenerative joint disease... [75].</p> <p>- A large number of participants said joint-related problems prevented them from engaging in PA. The loose term "arthritis" was used to describe weight-bearing large joint problems. "Breathing problems" was described as another reason for not engaging in PA [159].</p> <p>- ...individuals with coronary artery disease might avoid physical activity for fear of hurting themselves [143].</p>

Continued on next page

Barrier / Clas.	Prevented PAs	Source Quotation
Bad weather condition (cold, hot, icy, raining and wet)	Outdoor activities (e.g. walking, cycling and jogging)	<ul style="list-style-type: none"> - Weather is particularly cold..., making it difficult to exercise outdoors [240]. - For some of the participants, the weather was a barrier to physical activity. One participant said, "Right now the sun is hitting hard. ... one can't do it [172]. - I'm a bit reluctant especially in wet weather and icy weather... there's that many dips in footpaths now [34]. - ...weather was an important factor in PA behavior: "I'm kind of a sunshine walker. When it's raining, I'm not so good at it".. [29]. - ...hot summer climate (30–50 degree Celsius), which restricts outdoor activities like walking, cycling and jogging [25].
Lack of safety (crime or gang activity, poor street lighting, lack of neighbourhood safety, lack of safe places, unleashed dogs and unsafe walking).	Outdoor activities (e.g. walking)	<ul style="list-style-type: none"> - For some participants...poor street lighting, gang activity...were environmental barriers [172]. - ...crime is an important barrier to physical activity and functional ability in older populations [70]. - Another participant explained...There are no streetlights [172]. - ...many of the barriers..., including issues of..., neighbourhood safety,... [29]. - Street safety-women felt unsafe to walk alone in the city [25]. - ...Those dogs have scared me when going for walks with my boy in the stroller [172]. - African American focus groups reported numerous barriers, including unsafe walking areas [75].

Continued on next page

Barrier / Clas.	Prevented PAs	Source Quotation
Lack of facility or equipment (lack of bike paths, gyms, parks and swimming pool, and walking trails, and access to exercise places).	Different physical activities	<ul style="list-style-type: none"> - Reporting that higher SES groups described supportive environments, such as worksite facilities, bike paths, walking trails, gyms, and swimming pools [22]. - Overall, there was a lack of...and affordable exercise facilities or outdoor spaces for activities (e.g., parks) [25]. - Another participant explained, "There are no sidewalks. Not even an adequate park..." [172]. - Moreover, when examining environmental influences, women were more likely than men to report lack of available facilities...access to parks and facilities and perceived neighborhood environment (ie, sidewalks, enjoyable scenery, etc) [22]. - Other less commonly reported barriers also dealt with..., lack of access to exercise places and equipment [75].
Facilities and equipment cost	Different physical activities	<ul style="list-style-type: none"> - ...cost as barriers to physical activity [22]. - Also, among older adults, membership fees... often present insurmountable barriers to supervised programs in health clubs or recreational facilities [70].
Transportation problem (traffic, difficult parking, lack of transportation and car availability.	Different physical activities	<ul style="list-style-type: none"> - For some participants...transportation were environmental barriers [172]. - For some participants, traffic... and transportation were environmental barriers [172]. - Environmental factors... (e.g. ... difficult parking...) [29]. - ...lack of transportation often present insurmountable barriers to supervised programs in health clubs or recreational facilities [70]. - Excessive use of private cars reduced opportunities for physical activity (e.g., walking to bus stop, or walking to work) [25].

Continued on next page

Barrier / Clas.	Prevented PAs	Source Quotation
Lack of Knowledge (lack of physician advice, lack of professional guidance and lack of skills).	Different physical activities	<ul style="list-style-type: none"> - Other less commonly reported barriers...lack of physician advice [75]. - Most participants felt they had adequate knowledge and skills to perform regular physical activity [22]. - lack of guidance from a professional [29].
Lack of time (children restriction, family obligation, full time job, home responsibilities, take care of grand children).	Different physical activities (e.g. swimming and walking).	<ul style="list-style-type: none"> - The major barriers to physical activity was not having the time because of work and family obligations [172]. - Other less commonly reported barriers also dealt with social obligations (e.g., caring for children)... [75]. - Other participants agreed that taking care of their children or grand children [172]. - Maintaining a regular, adequate level of physical activity is challenging for adults who often have significant work and family responsibilities [108]. - ...women are more likely to have greater domestic responsibilities [22]. - I have to take care of my grandchildren... [240]. - There is no time because there is a lot of housework [12]. - sometimes children are at home so no time to exercise [12]. - Many respondents pointed to the difficulties of incorporating activities such as walking and swimming into what they regarded as extremely busy lives [142].

Continued on next page

Barrier / Clas.	Prevented PAs	Source Quotation
Old age	Different physical activities (e.g. running, squash).	<ul style="list-style-type: none"> - When I was 30 I used to go running sometimes, for about five years I would do regular exercise, play squash and go running... All of a sudden, it came back to me, this feeling for running, like a physical memory. At the moment I'm hoping I could do ten kilometres some day again [137]. - Age is also inversely related to physical activity rates [108].
Dislike, lack of enjoyment or lack of interest	Different physical activities.	<ul style="list-style-type: none"> - Other barriers to PA in general included dislike for or disinterest in exercise [29]. - Also enjoyment of sedentary activities can be a major barrier to increasing physical activity [143]. - Other prominent factors associated with general PA have included..., lack of time, enjoyment of PA [29]. - Other barriers to PA in general included dislike for or disinterest in exercise [29]. - ...Lack of enjoyment...[22].
Financial problem or lack of money	Different physical activities	<ul style="list-style-type: none"> - Individuals of low income reported lower self-efficacy than individuals of high income for participating in physical activity when tired [22]. - ...financial constraints... [108].
Laziness	Different physical activities.	<ul style="list-style-type: none"> - Barriers to regular physical activity included...laziness [136]. - After supper I'm so sleepy and get lazy...that's when it's really hard (to exercise) [240].
Preference for other activities	Different physical activities	<ul style="list-style-type: none"> - Other barriers to PA in general included...a preference for sedentary leisure-time activities [29].

Continued on next page

Barrier / Clas.	Prevented PAs	Source Quotation
Pain (hip pain, knee pain, leg pain, back pain and musculoskeletal pain, osteoarthritis and disability or injury).	Different physical activities (e.g. walk, bicycle and cricket)	<ul style="list-style-type: none"> - I can't ride a bicycle and walk as I used to because of problems with my legs, right [143]. - I have joint pain that keeps me from exercising [75]. - Individuals with physical complaints of..., hip/knee pain, or back pain, were more likely to report engaging in vigorous activity every week [203]. - "Because my legs start hurting...I have to sit periodically... [29]. - People with Type 2 diabetes report..., and individuals with coronary artery disease might avoid physical activity for fear of hurting themselves [143]. - Since osteoarthritis can be a barrier to physical activity [237]. - Many individuals had a physical disability/injury that sometimes made it difficult to participate in physical activity [22]. - When I found out I had diabetes, my legs started giving in, and I stopped playing cricket when I found out. I haven't played sports or anything since then [142].
Fatigue and tiredness	Different physical activities	<ul style="list-style-type: none"> - Fatigue and tiredness was a barrier reported in studies... [25]. - lack of energy and/or being tired are commonly reported barriers to physical activity among older adults [22]. - Of note, individuals with physical complaints of fatigue [203].
Lack of self-efficacy (lack of confidence)	Different physical activities	<ul style="list-style-type: none"> - ...and characteristics related to self-efficacy, such as confidence, expectations [29]. - They emphasized a lack of confidence about knowing their performance limit and believed that too much exercise could harm them, which made them feel insecure about a suitable level of physical activity [143].

Continued on next page

Barrier / Clas.	Prevented PAs	Source Quotation
Feeling depressed	Different physical activities	- ...Some participants cited..., depression,...as barriers [269].
Feeling stressed or anxious	Different physical activities	- ...anxiety....tend to be inversely associated with physical activity participation at any age (50–52) [70].
Lack of support and motivation (lack of family support and lack of partner support)	Different physical activities	- Low social support: “My family won’t encourage me; for example if I want to bring a walking machine they say there is no place to keep it” [12]. - Common barriers included...social support... [75]. - I do not have anyone to exercise with [75]. - Lack of motivation...were also mentioned as barriers to regular physical activity [22]. - Low motivation, lack of social support, competing demands...were the main barriers cited by the participants [12].
Cultural norms	Outdoor activities (e.g. walk)	- There are sociocultural norms that limit outdoor exercises, such as their families not allowing them to walk outside alone [12]. - ...due to gendered norms, social rules and cultural expectations, various barriers to physical activity appeared to pertain more to women than to men [142]. - Furthermore, whilst our respondents highlighted a dislike of going outdoors in bad weather in common with other cultural/ethnic groups [142].

Continued on next page

Barrier / Clas.	Prevented PAs	Source Quotation
Embarrassment	Different physical activities	<ul style="list-style-type: none"> - Socially, embarrassment, prioritizing domestic activities over PA...[159]. - The females said they were embarrassed and “uncomfortable” to engage in exercise in public areas [159]. - ...barrier and being ‘too shy/embarrassed,’...[157].
Overweight/obesity	Different physical activities	<ul style="list-style-type: none"> - Regardless of actual BMI, individuals who perceived themselves as overweight or obese were less active [108]. - ...while others report that feeling too obese to exercise or having various complications of obesity makes exercise difficult [143]. - Barriers to physical activity included...being overweight,... [269]. - According those earlier studies, most common barriers for regular exercise were....overweight,... [137]. - ...many patients with known T2D...often overweight or obese...do not achieve recommended levels of PA... [74]. - ...obesity is a barrier to increasing physical activity in Australian women [157].
Body image	Different physical activities	<ul style="list-style-type: none"> - ...poor body image has a negative relationship to activity levels [108]. - ...those who had a perceived body image in the largest third of pictured images had significantly lower activity levels in the domains of blocks walked per week and leisure time activity... [203].

Table A.3: Suggested physical activities (PAs) to limit identified barriers, with relevant source quotation

Barrier	Alternative Suitable PAs	Source Quotation
Hypoglycemia or low sugar level	Walking, gardening, general exercise, swimming, weight lifting, dancing, jumping, aerobics, yoga, biking, housework, yard work, shopping, car repair, taking elevators and stretching	<p>...walking is the preferred type of physical activity in our sample [172].</p> <p>- Participants said that physical activity programs should include walking, weight lifting, dancing, and swimming [172].</p> <p>- Most participants characterized physical activity as housework, yard work, dancing, shopping, and car repair. They also agreed that physical activity includes activities such as jumping, swimming, walking, aerobics, yoga, biking, and running. Most participants, however, preferred walking [172].</p> <p>- Moreover, it would be especially helpful to incorporate exercise into the everyday lives of older women (Haber, 2007) with diabetes such as walking instead of taking the bus or taking elevators and performing stretching indoors everyday [240].</p> <p>- Results from the qualitative telephone interviews revealed that walking was the most popular type of physical activity for all demographic groups [22].</p>

Continued on next page

Barrier	Alternative Suitable PAs	Source Quotation
Bad Weather condition	Indoor activities (conditioning Exercise and home activities)	- Consequently, men of higher income may have adequate resources to plan alternative strategies for physical activity when they cannot participate outdoors [22]. ...exercises that can be performed easily indoors should also be introduced as an alternative to outdoor exercise when it is too cold or hot [240].
Financial problems	Walking, housework and yard work	- In contrast, low income individuals were more likely to report walking for errands and activities of daily living including housework and yard work [22].
Lack of safety	Conditioning exercise and home activities	- Consequently, men of higher income may have adequate resources to plan alternative strategies for physical activity when they cannot participate outdoors [22]. - Unsafe neighborhood...PA options at home [29].
Lack of time	Walking, using stairs and stretching indoors	- Moreover, it would be especially helpful to incorporate exercise into the everyday lives...such as walking instead of taking the bus or taking elevators and performing stretching indoors everyday [240]. - Furthermore, walking can be easily incorporated into daily life.. [22].
No partner	Gym activities	It's nice to meet and talk to people; you meet other people at the gym [136].

Continued on next page

Barrier	Alternative Suitable PAs	Source Quotation
Old age	Climbing, gardening, walking, housework and playing with children	- Older people through short bouts of activities such as stair climbing, gardening, brisk walking, playing with children, or housework [70].
Cultural norms	Indoor activities	- Exercises that can be performed easily indoors should also be introduced as an alternative to outdoor exercise [240].
Osteoarthritis	Different physical activities	- Since osteoarthritis can be a barrier to physical activity, water-based physical activities, such as swimming, walking or running in a pool, or aquatic fitness classes have been encouraged for people with such comorbidities [70].
Difficult parking	Indoor activities	Difficult parking...PA options at home [29].

Appendix B

Additional Details of the Barrier Ontology Evaluation

B.1 Introduction

In Chapter 5, we applied two approaches for evaluation of the Barrier Ontology: the data-driven approach, and the competency questions (CQs) technique. The data-driven approach was applied to evaluate the terms or vocabularies of the Barrier Ontology. Due to the limits of the data-driven approach, CQs are used to evaluate the ontology’s ability to achieve its full purpose with quality and correctness. CQs are used to evaluate the consistency and integration among existing ontologies. Chapter 5 (Section 5.3.2) presents 4 templates of CQ, with two examples given for each (i.e. 8 CQs in total), ensuring the domain of the barrier is covered.

This appendix presents more details about both evaluation studies. For the data-driven approach, Section B.2 starts with a list of the 15 websites that were obtained via the search query ‘barriers to physical activity for Type 2 Diabetes (T2D) patients’ through the Google search engine¹. This is followed by presenting the complete table of the similarity between the Barrier Ontology and corpus. In addition, this appendix presents further details about the CQ technique. In Section B.3, 16 CQs are applied to evaluate the ability of the Barrier Ontology to achieve its requirements. An online survey is published principally to evaluate the return results of the CQs. The design of the survey is presented in Section B.3.1.

¹<https://www.google.com/>

B.2 Data-driven Approach

The data-driven approach is the second type of evaluation that is used to evaluate the Barrier Ontology. The Barrier Ontology was compared with existing data (the corpus) in the same domain. The terms or vocabularies of the Barrier Ontology are evaluated using the data-driven approach. We now give additional details about applying this approach to evaluate the Barrier Ontology.

Similarities between the Barrier Ontology and corpus: As part of the evaluation using the data-driven approach (Section 5.1), the Barrier Ontology was compared with the corpus. There was approximately 72% similarity (*F1-score*) between the Barrier Ontology and corpus. A full summary of these similarities is provided in Table B.1.

Table B.1: The complete similarity between terms of the Barrier Ontology and the corpus

Type of similarity	No.	Ontology	Corpus
Syntactical	1	Hypoglycaemia	Hypoglycaemia
	2	Heart disease	Heart disease
	3	Pain (injury, knee)	Pain (injury, knee)
	4	Lack of time	Lack of time
	5	Age	Age
	6	Fatigue	Fatigue
	7	Embarrassment	Embarrassment
	8	Anxious	Anxious
	9	Obesity	Obesity
Synonymous	10	Lack of safety	Safety (of your home)
	11	Cost of equipments	Cost of sport equipments
	12	Dislike	Disliking
	13	Feeling lazy	Laziness
	14	Body image	Body fat
	15	Cultural norms	cultural issues
	16	Lack of facilities	Lack of sport facilities
	17	Lack of self-efficacy	low perceived self-efficacy
	18	Lack of motivation	Motivation
	19	Weather condition	Climatic condition

Extraction of Text Corpus: A text corpus is extracted from the knowledge domain on which the Barrier Ontology is based. This corpus was generated from 15 website articles, after removing two duplicate sites. It is worthy to note that, these sites are completely exclusive of the articles we used in the systematic review (Section 4.3.2). The 15 crawled

websites are obtained through the Google search engine using combinations of keywords such as, ‘barrier’ or ‘obstacle’, and ‘physical activity’ or ‘exercise’. Table B.2 lists these 15 websites, including the associated Uniform Resource Locator (URL) and content type (i.e. article or web page).

Table B.2: Information about the websites crawled for information in generating the corpus.

No.	URL	Article / Web	Title
1	https://pmj.bmj.com/content/80/943/287	Article	Barriers to physical activity in patients with diabetes
2	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5749073/	Article	Exercise in patients with Type 2 diabetes: Facilitators and barriers - A qualitative study
3	https://academic.oup.com/qjmed/article/1e/106/7/635/1592102	Article	Barriers to exercise in obese patients with type 2 diabetes
4	https://clinicaltrials.gov/ct2/show/NCT01701570	Web page	Barriers to Physical Activity in People With Type 2 Diabetes (Exercise)
5	https://care.diabetesjournals.org/content/31/11/2108	Web page	Barriers to Physical Activity Among Patients With Type 1 Diabetes
6	https://professional.diabetes.org/abstract/perceived-barriers-exercise-patients-type-2-diabetes-mellitus	Article	Perceived Barriers to Exercise in Patients with Type-2 Diabetes Mellitus
7	https://journals.sagepub.com/doi/10.1177/0145721713492565	Article	Barriers to Physical Activity in People With Type 2 Diabetes Enrolled in a Workplace Diabetes Disease Management Program
8	https://onlinelibrary.wiley.com/doi/full/10.1111/hsc.12263	Article	A cross-sectional study of barriers to physical activity among overweight and obese patients with type 2 diabetes in Iran
9	https://onlinelibrary.wiley.com/doi/full/10.1002/pdi.2136	Article	Exercise and type 1 diabetes: overcoming the barriers

Continued on next page

Continued from previous page

No.	URL	Article / Web	Title
10	https://www.hindawi.com/journals/isrn/2017/1273084/	Article	Barriers to Diet and Exercise among Nepalese Type 2 Diabetic Patients
11	https://livrepository.liverpool.ac.uk/3008268/1/p16-alfaifiAFL.pdf	Article	Towards an Ontology to Identify Barriers to Physical Activity for Type 2 Diabetes
12	https://www.physiotherapyjournal.com/article/S0031-9406(15)00521-0/fulltext	Article	Identification of barriers for adherence to exercise in type 2 diabetes mellitus—a cross sectional observational study
13	https://en.wikibooks.org/wiki/Exercise_as_it_relates_to_Disease/Barriersto_Physical_Activity_Among_Patients_With_Type_1_Diabetes	Web page	Exercise as it relates to Disease/Barriers to Physical Activity Among Patients With Type 1 Diabetes
14	https://bjgp.org/content/60/577/570	Article	Type 2 diabetes and dog walking: patients' longitudinal perspectives about implementing and sustaining physical activity
15	https://www.researchgate.net/publication/265859408_Attitudes_and_Barriers_to_Exercise_in_Adults_with_Type_1_Diabetes_T1DM_and_How_Best_to_Address_Them:_A_Qualitative_Study	Article	Attitudes and Barriers to Exercise in Adults with Type 1 Diabetes (T1DM) and How Best to Address Them: A Qualitative Study

B.3 Competency Questions Technique

We present 16 competency questions (CQs) in total, each fitting into one of four CQ templates detailed in Section 5.3.2. Each template is responsible for verifying that a specific requirement (Section 4.2) of the Barrier Ontology has been met. These CQs were (1): transformed into formal queries (SPARQL) and run using Protégé (Section 3.6) to retrieve their answers from the ontology and (2): used for creating a survey questionnaire in which domain experts (physicians, clinicians and academics, etc.) provide feedback in the form of answers to the CQs. Important to note that, the minimum threshold of votes (survey responses) required to consider an answer is set at 10 votes (approximately 33% of 30 total participants), implying that answers having fewer responses are excluded from the analysis. This section supplements the content covered in CQ evaluation section of Chapter 5 (Section 5.3). It is organised as follows: a narration of the design of the online survey questionnaire used to elicit knowledge in form of answers to CQs from experts, followed by template CQs, their respective queries and results (BO), then survey results (SR) to the CQs. Finally a table for each template that reveals how BO answers varied from SR answers. The intrinsic analytic comparison discussion between BO and SR is covered under the discussion sections of each template in Chapter 5.

B.3.1 Survey Design and Content

The online survey questionnaire is a mixture of both structured (fixed responses) and non-structured (open-ended) questions. Each question was constructed on the basis of at least one functional requirement of the ontology and to supplement one or more CQs used in evaluation of the ontology. Survey questions, with their corresponding CQ and functional requirements, are presented in Table B.3.

The aim of following this pattern in creating the survey was to uphold the coherence of the propositions, results and evaluation in the thesis. For instance, given the above scenario, we shall determine the extent to which the results obtained by the ontology matched the survey-oriented results. Table B.3 describes how survey questions were developed with respect to functional requirements and CQs. It additionally consists of the particular relation responsible for establishing the links necessary to retrieve answers to the CQ. In other words, the quoted relation forms the underlying links between the ontology classes to be checked when answering the CQ. In addition to the various question formats used i.e. multiple choice, fixed response and open-ended, the questions were organised in an order

that could be sequentially comprehended by participants, thereby starting with questions querying the participants’ occupation, followed by subject-related queries, and finishing with survey evaluation feedback questions. Besides the participant’s profession and qualification, questions requesting personal information such as name, age, gender, background were excluded from the survey.

An ethics approval from the University of Liverpool, Department of Ethics², has been obtained to conduct the online survey titled “Ontology to Support Behaviour Change Intervention: Barriers to Activity”, which we built and published using a web-based platform (AllCounted³).

We utilise purposive sampling, a non-probability sampling technique to identify participants or the target group for the online survey. While probability sampling gathers participants randomly, non-probability sampling is more discriminative in its selection process. Purposive sampling is a criterion based non-probability sampling where researchers select an information-rich data source for participation [251]. This survey targets qualified professionals aged 18+ in the following occupations: clinical psychologist, psychiatrist, endocrinologist, diabetes specialist, other medical doctor, mental health nurses, other nurses, allied health care professional and academic researchers. Other people included in the target group were people who had keen interest in T2D. A comprehensive review and analysis of the survey results is given in Chapter 5 (Section 5.3.2).

Table B.3: A table of various CQs of each template, with their corresponding survey questions, target relations, and the ontological functional requirements they satisfy. Additional comments narrating how survey questions are derived from CQs are included as Meta-data.

Functional requirements	require-	To classify the identified barriers into one of the six main categories (Health, Environmental barriers, etc.).
Template 1		CQ 1.1: What barriers are classified as ‘ <i>Health</i> ’ barrier? CQ 1.2 What barriers are classified as ‘ <i>Environmental</i> ’ barrier?

Continued on next page

²<https://www.liverpool.ac.uk/research-integrity/research-ethics/>

³<https://www.allcounted.com/privacy?uid=sf15f2zr05vp1>

Table B.3 – *Continued from previous page*

Survey question	For each barrier (listed on the left) please select the category (listed along the top) that you believe is the most appropriate. You can select a category by clicking or tapping on the radio button. You can select only one category for each barrier and leave blank any barrier for which you are unsure.
Meta-data	Each CQ under Template 1 precisely requires barriers that belong to a single barrier type or category. For example, CQ 1.1. requires barriers belonging to the “Health Barriers” type, CQ 1.2 requires barriers of type “Environmental barriers” etc. We design a single survey question in which participants are required to classify barriers into barrier types or categories.
Target relation	subClassOf
Functional requirement	To identify the barriers, based on user characteristics, that may limit the user from performance of physical activity.
Template 2	CQ 2.1: What is the expected barrier for a male patient who is 37 years old, has a full-time job, lives with his wife and two children, and complains about not having friends to play sports with, and no personal support for performing physical activity?
Survey question	What barriers are experienced by patient in scenario below? "Tony is aged 37, married, and a full-time office worker who usually works overtime at the weekend. He lives with his wife and two children, aged 2 and 5. He is in good health and does not suffer from any diseases. Tony complains about not having friends to play sports with, nor any personal support for performing physical activity.

Continued on next page

Table B.3 – *Continued from previous page*

Meta-data	The CQs in Template 2 require barriers that impede a patient’s engagement in physical activity, given various characteristics or attributes of a patient. For each CQ, a survey question (in the form of a scenario) is designed to include all the key patient characteristics in the CQ, and thereby ask survey participants to respond indicating which barriers they expect the patient to experience. For example, in the CQ 3.1 scenario presented above participants respond with potential barriers to the patient’s physical activity engagement. Two CQs and corresponding scenarios for Template 3 are presented and analysed in Chapter 5, and 2 others in this appendix.
Target relation	<i>BarrierSign</i> sub-property relations (e.g. <i>age</i> , <i>hasFullTimeJob</i> , <i>isMarriedWithChildren</i> , <i>isSocial</i>)
Functional requirement	To recognise barriers that <i>prevent</i> performance of a specific type of activity.
Template 3	CQ 3.1: What are the barriers that prevent or limit people from performing ‘ <i>Football</i> ’, despite being interested in this activity? CQ 3.2: What are the barriers that prevent or limit people from performing ‘ <i>Swimming</i> ’, despite being interested in this activity?
Survey question	For each barrier (listed on the left), please select which activity types (listed along the top) you believe the barrier could <i>prevent</i> people to engage into. For each barrier, you may select multiple answers. Leave blank the boxes for any barrier for which you are unsure.

Continued on next page

Table B.3 – *Continued from previous page*

Meta-data	The 2 CQs require barriers that prevent a single physical activity each i.e. football (CQ 3.1), swimming (CQ 3.2) etc. However, we design the corresponding survey question to probe for barriers that categorically prevent several physical activities presented in CQs within the template, i.e. the survey question requires barriers that prevent aerobic athletic sports (e.g. football for CQ 3.1) and conditioning exercises (e.g. swimming for CQ 3.2) etc.
Target relation	<i>prevents</i> relation
Functional requirement	To suggest activities to limit identified barriers.
Template 4	CQ 4.1: What physical activity is suggested for the bad weather condition (<i>‘Raining’</i>) barrier? CQ 4.2: What physical activity is suggested for the lack of time (<i>‘Home responsibilities’</i>) barrier?
Survey question	For each barrier (listed on the left), please select which activity types (listed along the top) you believe could be ‘Suggested’ to people experiencing the barrier, to mitigate the barrier’s effect. For each barrier, you may select multiple answers. Leave blank the boxes for any barrier for which you are unsure.

Continued on next page

Table B.3 – *Continued from previous page*

Meta-data	Similar to Template 3, the 2 CQs above require activities that are suitable for a single barrier each, bad weather condition (Raining) in the case of CQ 4.1, lack of time (Home responsibilities) in CQ 4.2 etc. The corresponding survey question (for this templates CQs) is designed to probe for suitable activities suggested for people experiencing several barriers (including those stated in the distinguished CQs within the template) and many other barriers in the ontology. Contextually, the survey question requests participants suggest activities for patients experiencing bad weather condition (CQ 4.1), lack of time (CQ 4.2) etc. We consider the subset of the answers to the survey question for each CQ. For instance, for CQ 4.1, we concentrate on activities suggested for bad weather condition.
Target relation	<i>isSuggestedFor</i> relation

B.3.2 Templates of Competency Questions

B.3.2.1 Template 1

Below are the SPARQL queries corresponding to each CQ under template 1:

CQ 1.1: What barriers are classified as ‘*Health*’ barrier?

```

1 Query 1.1:
2 SELECT distinct ?Barrier
3     WHERE {
4         {?Barrier rdfs:subClassOf :HealthBarrier}
5         UNION
6         {?Barrier rdf:type :HealthBarrier}
7     }
```

```

Query 1.2 results:
- MedicalCondition          - Hypoglycemia
```

CQ 1.2: What barriers are classified as ‘*Environmental*’ barrier?

```

1 Query 1.2:
```


Survey Results (SR) and comparison with Barrier Ontology (BO), for questions corresponding to Template 1. Figure B.1 visualises the classification for the granular barriers into their respective categories, as per survey feedback collected from domain experts for CQs under Template 1. 87% of the participants classified ‘hypoglycemia’ as a ‘Health barrier’, and 47% classified ‘fatigue’ as a ‘Physical barrier’. Furthermore, 80% of participants claim that ‘financial problem’ is a ‘Personal barrier’ and 97% indicate ‘weather condition’ (cold, hot and wet) as ‘Environmental barrier’. Some barriers had multiple classifications i.e. according to the results, they were suited to belong to more than one barrier category. An example of this is where ‘fatigue’ was classified as both a ‘Health’ and ‘Physical’ barrier, receiving 33% and 47% votes respectively.

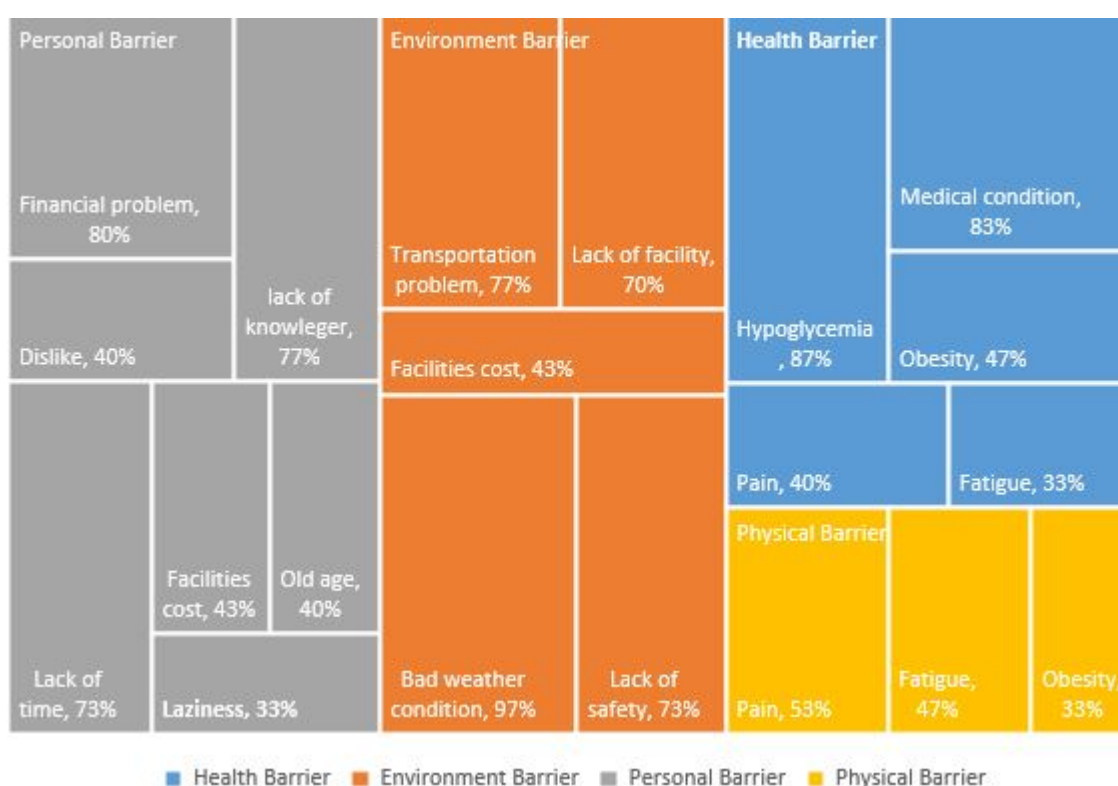


Figure B.1: A tree map depicting how domain experts classified barriers for CQs under Template 1. Each barrier classification was confirmed by a minimum of 33% of the total votes.

Table B.4 reveals answers corresponding to CQs 1.3 and 1.4 under Template 1 (answers corresponding to CQs 1.1 and 1.2 are presented in Chapter 5), in terms of classification

choices made by the experts (SR) along with classifications retrieved from the BO, including a percentage similarity value (PS) indicating how comparable BO was to SR. For example, a PS of 67% was obtained for CQ 1.4 resulting from the fact that, out of the combined total of barriers classified, BO classified only 2 of these as Physical, whereas SR classified all 3 as Physical, hence the 67% PS value.

Table B.4: Classifications assigned to barriers according to the Barrier Ontology (BO) results and Survey Result (SR) for CQs 1.3 and 1.4 under Template 1. ✓ implies a barrier is identified as belonging to the corresponding Barrier Type, whereas X is the opposite. A match (Percentage Similarity (PS)) between the BO and SR column indicates that the ontology retrieved answers identical to those provided by experts in the survey.

CQ No.	Barrier Type	Barriers	BO	SR	PS %
CQ 1.3	Personal Barrier	Lack of knowledge	✓	✓	75%
		Lack of time	✓	✓	
		Old age	✓	✓	
		Dislike or lack of interest	✓	✓	
		Financial problem	✓	✓	
		Laziness	✓	✓	
		Preference for other activities	✓	X	
		Facilities and equipment cost	X	✓	
CQ 1.4	Physical Barrier	Physical pain	✓	✓	67%
		Fatigue and tiredness	✓	✓	
		Obesity or overweight	X	✓	

B.3.2.2 Template 2

Below are the SPARQL queries corresponding to each CQ under Template 2:

CQ 2.1: What is the expected barrier for a male patient who is 37 years old, has a full-time job, lives with his wife and two children, and complains about not having friends to play sports with, and no personal support for performing physical activity?

```

1 Query 2.1:
2 SELECT distinct ?barriers
3 WHERE {
4     values ?dtps { :under60 :isMarriedWithChildren
5                   :hasFullTimeJob :isSocial }.
6     values ?labels { "yes"^^xsd: "yes"^^xsd: "yes"^^xsd: "no"^^xsd: }.
7     ?dtps rdfs:seeAlso ?barriers.
8     ?dtps rdfs:comment ?labels
9 }
```

Query 2.1 results:

- LackOfTime - LackOfSupportAndMotivation

CQ 2.2: What is the expected barrier for a person who has a full-time job, a minor injury, and lives in a very crowded area with traffic issues? Their nearest sports centre or gym and walking trails are more than an hour away by bus?

```

1 Query 2.2:
2 SELECT distinct ?barriers
3 WHERE {
4     values ?dtps { :under60 :hasFullTimeJob :hasPain
5                   :traffic :trainingFacility }.
6     values ?labels { "yes"^^xsd: "yes"^^xsd: "yes"^^xsd: "poor"^^xsd:
7                     "inaccessible"^^xsd: }.
8     ?dtps rdfs:seeAlso ?barriers.
9     ?dtps rdfs:comment ?labels
10 }
```

Query 2.2 results:

- LackOfTime - TransportationProblem
- PhysicalPain - LackOfFacilityAndEquipment

CQ 2.3: What is the expected barrier for a 70-year-old person who lives in an area with very changeable weather where it is often raining and cold? They suffer from occasional back pain.

```

1 Query 2.3:
2 SELECT distinct ?barriers
3 WHERE {
4     values ?dtps { :isOver60 :weather :hasBackPain }.
5     values ?labels { "yes"^^xsd: "cold_and_raining"^^xsd: "yes"^^xsd: }.
6     ?dtps rdfs:seeAlso ?barriers.
7     ?dtps rdfs:comment ?labels
8 }
```

Query 2.3 results:

- BadWeatherCondition - PhysicalPain
- OldAge

CQ 2.4: What is the expected barrier for a 45-year-old person who lives in an area that lacks affordable gym facilities and suffers from asthma? They spend lots of time watching

TV and browsing the internet.

```

1 Query 2.4:
2 SELECT distinct ?barriers
3 WHERE {
4     values ?dtps { :isUnder60 :hasOtherDiseases
5                   :nonPhysicalActivity :trainingFacility}
6     values ?labels { "yes"^^xsd: "Asthma"^^xsd: "inaccessible"^^xsd:
7                      "Television_or_Internet_browsing"^^xsd:}.
8     ?dtps rdfs:seeAlso ?barriers.
9     ?dtps rdfs:comment ?labels
10  }
```

Query 2.4 results:

```

- PreferenceForOtherActivities      - MedicalCondition
- LackOfFacilityAndEquipment
```

Survey Results (SR) and comparison with Barrier Ontology (BO), for questions corresponding to Template 2: Table B.5 presents results to all CQs under Template 2. Each value in the table represents the percentage volume of votes that a barrier received, for qualifying as an activity preventing barrier with respect to the different CQs. Each CQ was presented as a scenario in the survey, i.e. CQ 2.1 corresponds to scenario 1, CQ 2.2 corresponds to scenario 2, and so on. Generally, the percentage volume of votes was low, ranging between 33% and 57%. Nevertheless, this volume satisfied our requirement of filtering out answers from the survey, i.e. an answer was only considered if it had received at least 33% of the total votes (as detailed in Section 5.3.2). With the exception of ‘lack of time’ and ‘lack of facility’, all barriers were voted for in only one of the four presented scenarios. For example, ‘lack of support’ received votes only in scenario 1 (47%), ‘pain(injury)’ received votes only in scenario 3 (38%), ‘transportation problem’ received votes only in scenario 2 (20%), and so on. It is also observed that the survey participants voted for at least 3 barriers as activity preventing barriers for the patients described in scenarios 2 and 3, returning ‘lack of time’, ‘transportation problem’ and ‘lack of facility’ in scenario 2, and returning ‘pain’, ‘weather’ and ‘old age’ in scenario 3. This is in contrast to scenarios 1 and 4, where there was only two barriers returned by the survey participants.

Table B.6 reveals how BO and SR varied for the Template 2 CQs. There was a perfect match between activity preventing barriers in BO and SR for scenario 3, resulting in

Table B.5: Percentage volume of votes by domain experts, for activity preventing barriers for patients described in scenarios derived from Template 2 CQs.

Barrier	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Lack of time	57%	33%	-	-
Lack of support/motivation	47%	-	-	-
Lack of facility	-	53%	-	40%
Transportation problem	-	50%	-	-
Bad weather	-	-	52%	-
Old age	-	-	38%	-
Pain (e.g. injury)	-	-	38%	-
Medical condition (asthma)	-	-	-	37%

a 100% PS. That is to say, both BO and SR returned ‘bad weather condition’, ‘old age’ and ‘physical pain’ as barriers that can hinder the patient described in scenario 3 (CQ 2.3).

Table B.6: Potential barriers according to the Barrier Ontology (BO) and Survey Results (SR) for scenarios 3 and 4 under Template 2. ✓implies a barrier is a hindrance to patient in response to the CQs, whereas X is the opposite.

Barriers	Scenario 3 (CQ 2.3)		Scenario 4 (CQ 2.4)	
	BO	SR	BO	SR
Weather condition	✓	✓	-	-
Old age	✓	✓	-	-
Physical pain	✓	✓	-	-
Preference for other activities	-	-	✓	X
Medical condition (asthma)	-	-	✓	✓
Lack of facility	-	-	✓	✓
Percentage Similarity (PS)	100%		67%	

B.3.2.3 Template 3

Below are the SPARQL queries corresponding to each CQ under Template 3:

CQ 3.1: What are the barriers that prevent or limit people from performing ‘Football’, despite being interested in this activity?

```

1 Query 3.1:
2 SELECT distinct ?barrier
3     WHERE { ?barrier :prevents :Football.
4             :prevents rdf:type owl:ObjectProperty }
```

Query 3.1 results:

- LackOfSavePlaces	- LackOfSkills	- OldAge
- PoorAccessToExercisePlaces	- FullTimeJob	- Asthma
- LackOfTransportation	- ColdWeather	- KneePain
- PreferenceForOtherActivities	- DislikeActivity	- Obesity
- Hypoglycemia	- LackOfConfidence	- BodyImage
- FeelingStressedOrAnxious	- CulturalNorms	- Embarrassment
- FatigueAndTiredness	- FeelingDepressed	- Laziness

CQ 3.2: What are the barriers that prevent or limit people from performing ‘Swimming’, despite being interested in this activity?

```

1 Query 3.2:
2 SELECT distinct ?barrier
3     WHERE {
4         ?barrier :prevents :Swimming.
5         :prevents rdf:type owl:ObjectProperty
6     }

```

Query 3.2 results:

- PreferenceForOtherActivities	- Laziness	- OldAge
- LackOfSwimmingPool	- HomeResponsibilities	- HeartRate
- LackOfTransportation	- LackOfSavePlaces	- BackPain
- FacilitiesAndEquipmentCost	- DislikeActivity	- Obesity
- LackOfProfessionalGuidance	- LackOfConfidence	- BodyImage
- FeelingStressedOrAnxious	- FinancialProblem	- Embarrassment
- FatigueAndTiredness	- FeelingDepressed	

CQ 3.3: What are the barriers that prevent or limit people from performing ‘Yard work’, despite being interested in this activity?

```

1 Query 3.3:
2 SELECT distinct ?barrier
3     WHERE {
4         ?barrier :prevents :YardWork.
5         :prevents rdf:type owl:ObjectProperty
6     }

```

Query 3.3 results:

- FeelingStressOrAnxious	- FeelingDepressed	- Laziness
- PreferenceForOtherActivities	- DisabilityOrInjury	- HeartRate

- FatigueAndTiredness	- HomeResponsibilities	- DislikeActivity
-----------------------	------------------------	-------------------

CQ 3.4: What are the barriers that prevent or limit people from performing ‘*Running*’, despite being interested in this activity?

```

1 Query 3.4:
2 SELECT distinct ?barrier
3     WHERE {
4         ?barrier :prevents :Running.
5         :prevents rdf:type owl:ObjectProperty
6     }

```

Query 3.4 results:

- PreferenceForOtherActivities	- OldAge	- LackOfPark
- FamilyObligation	- Hypoglycemia	- HeartRate
- DisabilityOrInjury	- RainyWeather	- Laziness
- LackOfSavePlaces	- DislikeActivity	- Obesity
- FeelingDepressed	- LackOfConfidence	- BodyImage
- FeelingStressedOrAnxious	- CulturalNorms	- Embarrassment
- FatigueAndTiredness	- LackOfPartnerSupport	

Survey Results (SR) and comparison with Barrier Ontology (BO), for questions corresponding to Template 3: Table B.7 presents results for all CQs under Template 3. Each value in the table represents the percentage of votes a barrier received for qualifying as a barrier that prevents a specific activity quoted in the CQs. It is vividly clear that besides ‘Yard work’, all other activities have got several barriers, with ‘Football’ topping the list having 23 barriers in total. It is observed that a few barriers received a low percentage volume of votes, such as ‘lack of safety’ (33%) and ‘facilities and equipment cost’ (43%). Nevertheless, these barriers are considered because they satisfy the requirement set to filter out answers from the survey (answers are considered provided they receive 33% of the votes).

Tables B.8 and B.9 reveal the variation between BO and SR for CQ 3.3 and CQ 3.4 respectively. BO and SR concurred on 9 barriers out of 11, resulting in a PS of 82% for CQ 3.3. Additionally BO and SR concurred on 19 of the 22 different barriers selected, resulting into a PS value of 86% for CQ 3.4.

For all CQs under this template, there is a significant similarity between SR and BO. The lowest PS value is 82% for ‘Yard work’ activity.

Table B.9: Potential barriers to *Running* activity according to Barrier Ontology (BO) and Survey results (SR)

No	Barrier	BO	SR	No	Barrier	BO	SR
1	Hypoglycemia	✓	✓	12	Heart rate	✓	✓
2	Preference for other activities	✓	✓	13	Laziness	✓	✓
3	Fatigue and tiredness	✓	✓	14	Dislike activity	✓	✓
4	Lack of save places	✓	✓	15	Rainy weather	✓	✓
5	Lack of confidence	✓	✓	16	Lack of park	✓	✓
6	Lack of partner support	✓	✓	17	Feeling depressed	✓	✓
7	Family obligation	✓	✓	18	Cultural norms	✓	✓
8	Transportation problem	X	✓	19	Obesity	✓	✓
9	Lack of Knowledge	X	✓	20	Body image	✓	✓
10	Embarrassment	✓	✓	21	Old age	✓	X
11	Feeling stressed or anxious	✓	✓	22	Disability or injury	✓	✓
Percentage Similarity (PS) = 86.36%							

B.3.2.4 Template 4

Below are the SPARQL queries corresponding to each CQ under Template 4:

CQ 4.1: What physical activity is suggested for the bad weather condition (*‘Raining’*) barrier?

```

1 Query 4.1:
2 SELECT distinct ?PhysicalActivity
3     WHERE {
4         ?PhysicalActivity :isSuggestedFor :RainyWeather.
5         :isSuggestedFor rdf:type owl:ObjectProperty
6     }
```

```

Query 4.1 results:
- HomeCleaning          - Yoga
- WeightLifting         - RopeSkipping
- Swimming              - WalkingUpstairs
- Stretching            - PushUps
- HouseholdTasks
```

CQ 4.2: What physical activity is suggested for the lack of time (*‘Home responsibilities’*) barrier?

```

1 Query 4.2:
2 SELECT distinct ?PhysicalActivity
```

```

3      WHERE {
4          ?PhysicalActivity :isSuggestedFor :HomeResponsibilities.
5          :isSuggestedFor rdf:type owl:ObjectProperty
6      }

```

Query 4.2 results:

- HouseholdTasks	- RopeSkipping
- Yoga	- WalkingUpstairs
- HomeCleaning	- PushUps
- Walking	

CQ 4.3: What physical activity is suggested for the ‘*Financial problem*’ barrier?

```

1 Query 4.3:
2 SELECT distinct ?PhysicalActivity
3     WHERE {
4         ?PhysicalActivity :isSuggestedFor :FinancialProblem.
5         :isSuggestedFor rdf:type owl:ObjectProperty
6     }

```

Query 4.3 results:

- LawnAndGardenActivity	- YardWork
- Walking	- RopeSkipping
- HouseholdTasks	- WalkingUpstairs
- HomeCleaning	- PushUps

CQ 4.4: What physical activity is suggested for the ‘*Cultural norms*’ barrier?

```

1 Query 4.4:
2 SELECT distinct ?PhysicalActivity
3     WHERE {
4         ?PhysicalActivity :isSuggestedFor :CulturalNorms.
5         :isSuggestedFor rdf:type owl:ObjectProperty
6     }

```

Query 4.4 results:

- HomeCleaning	- Yoga
- WeightLifting	- PushUps
- Swimming	- HouseholdTasks
- Stretching	

Survey Results (SR) and comparison with Barrier Ontology (BO), for questions corresponding to Template 4: Figure B.2 uses a grouped bar chart to visualise the suggested activities for patients struggling with different barriers in Template 4 CQs. Home activities are suggested as suitable activities for patients struggling with all four barriers presented in the four respective CQs. Occupational activities are the second most popular activities suggested for patients struggling with three barriers in CQs 3.2, 3.3 and 3.4. Conditioning exercise was the least popular suggested activity, nominated as a suitable activity only for the ‘bad weather condition’ barrier in CQ 4.1.

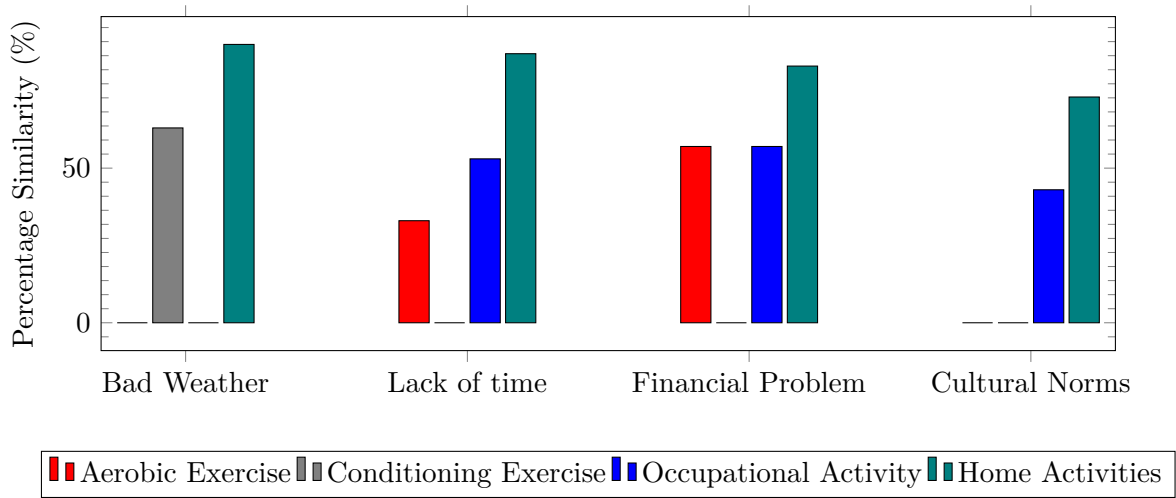


Figure B.2: Bar chart indicating the category of activities that domain experts nominated as suitable for T2D patients who experience four specific barriers: ‘bad weather condition’, ‘lack of time’, ‘financial problem’ and ‘cultural norms

Table B.10 compares suggested activities for the financial problem barrier (CQ 4.3) and cultural norms barrier (CQ 4.4).

Table B.10: Suggested activities according to the Barrier Ontology (BO) and Survey Results (SR) to limit financial problem barrier (CQ 4.3) and cultural norms barrier (CQ 4.4) under Template 4. ✓implies an activity is a suggestion to patient in response to the CQs, whereas X is the opposite.

	Financial problem (CQ 4.3)		Cultural norms (CQ 4.4)	
Physical activities	BO	SR	BO	SR
Aerobic exercises	✓	✓	-	-
Conditioning exercises	-	-	✓	X
Occupational activities	✓	✓	X	✓
Home activities	✓	✓	✓	✓
Percentage Similarity	100%		33%	